



Universidad Pública de Navarra  
Nafarroako Unibertsitate Publikoa

**ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA  
AGRONÓMICA Y BIOCENCIAS**

***NEKAZARITZAKO INGENIARITZAKO ETA BIOZIENTZIETAKO  
GOI MAILAKO ESKOLA TEKNIKOA***

**RELLENO DE HUECOS EN SERIES TEMPORALES DE ÍNDICES  
SENTINEL-2 EN CUBIERTAS AGRÍCOLAS CON DATOS  
PROCEDENTES DE SENTINEL-1**

presentado por

ITSASO ARANGUREN ERICE (*e*)k

*Aurkeztua*

**MASTER UNIVERSITARIO EN SISTEMAS DE INFORMACIÓN GEOGRÁFICA Y  
TELEDETECCIÓN**

***UNIBERTSITATE MASTERRA INFORMAZIO GEOGRAFIKOKO SISTEMETAN ETA  
TELEDETEKZIOAN***



Septiembre, 2021 / 2021, iraila



## **AGRADECIMIENTOS**

*Quiero agradecer, en primer lugar, a mi tutor Jesús Álvarez Mozos y directora Lourdes Albizua Huarte por vuestra dedicación a lo largo de todo el trabajo, sin vosotros no habría sido posible sacarlo adelante. Agradecer también a María Arias por haberme proporcionado las imágenes ya procesadas a partir de las cuales se ha realizado el trabajo. De igual forma, agradecer a la Universidad Pública de Navarra por la concesión de la Ayuda de Iniciación a la Investigación, habiendo realizado el trabajo en el ámbito del Instituto de Investigación ISFOOD (Institute for Innovation & Sustainable Development in Food Chain).*

*Finalmente, agradecer a mi familia y amigas el apoyo incondicional proporcionado siempre.*





## RESUMEN

La teledetección ofrece la posibilidad de realizar un seguimiento del desarrollo de los cultivos mediante la adquisición de imágenes de forma repetida por sensores alojados en plataformas satelitales. El objetivo del presente trabajo fin de master es rellenar la falta de información en las series temporales de índices Sentinel-2 (sensor óptico-multiespectral) causadas por las nubes, con datos procedentes de Sentinel-1 (sensor radar cuya funcionalidad no se ve limitada por la nubosidad). Para ello, se calcularon y evaluaron índices ópticos y radar que ofrecieran una mayor sensibilidad al desarrollo de siete cultivos representativos de una zona de estudio en Navarra. Posteriormente, se realizó un análisis estadístico y evaluó la correlación entre los índices calculados. Finalmente, se estudiaron y evaluaron las técnicas de relleno aplicadas. Los resultados muestran que podría ser factible rellenar la falta de información en series temporales de NDVI a través de datos de índices de Sentinel-1.

## Palabras clave

Índice de vegetación, relleno de huecos, Sentinel-1, Sentinel-2, series temporales, teledetección.

---

## ABSTRACT

Remote sensing offers the possibility to track crop development by repeatedly acquiring images by sensors housed on satellite platforms. The objective of this work is to fill in the gaps of information in the vegetation indices time series of Sentinel-2 (optical-multispectral sensor) caused by clouds, with data from Sentinel-1 (radar sensor whose functionality is not seen limited by cloud cover). First, optical and radar indices that offered greater sensitivity to the development of 7 representative crops of a study area in Navarra, were calculated and evaluated. Then, a statistical analysis was performed and the correlation between the calculated indices was evaluated. Finally, the applied gap filling techniques were studied and evaluated. Results show that it might be feasible to fill the gaps in NDVI time series using Sentinel-1 index data.

## Key words

Gap filling, remote sensing, Sentinel-1, Sentinel-2, time series, vegetation index.



## ÍNDICE

1. INTRODUCCIÓN .....	1
1.1. Programa Copernicus .....	1
1.2. Teledetección y agricultura .....	4
1.3. Técnicas de relleno de huecos .....	7
2. OBJETIVOS .....	10
3. MATERIAL .....	11
3.1. Descripción de la zona de estudio .....	11
3.2. Datos de partida .....	12
4. METODOLOGÍA .....	18
4.1. Análisis preliminar .....	19
4.2. Aplicación de las técnicas de relleno de huecos .....	24
5. RESULTADOS Y DISCUSIÓN .....	28
5.1. Selección del cálculo de índices a nivel pixel o a nivel parcela .....	28
5.2. Selección de pasada de Sentinel-1 .....	30
5.3. Selección de índices .....	30
5.4. Selección de unidades .....	31
5.5. Selección de ventanas temporales .....	33
5.6. Selección de suavizado de curvas .....	34
5.7. Análisis del comportamiento reflexivo y retrodispersivo de los cultivos .....	36
5.8. Evaluación del impacto de los huecos .....	42
5.9. Evaluación de las técnicas de relleno de huecos .....	46
5.10. Comparación entre técnicas de relleno .....	56
6. CONCLUSIONES .....	60
7. BIBLIOGRAFÍA .....	63
ANEXOS .....	68

## ÍNDICE DE FIGURAS

<b>Figura 1.</b> Comparación de las bandas espectrales de Sentinel-2 y Landsat 8 y 9. <i>Fuente:</i> Poursanidis & Chrysoulakis (2017) .....	3
<b>Figura 2.</b> Curva de respuesta espectral para distintas coberturas. <i>Fuente:</i> (Chuvieco, 2006) .....	5
<b>Figura 3.</b> Situación de la zona de estudio. ....	11
<b>Figura 4.</b> Huella satelital (zona sombreada en rojo) de la pasada ASC_103 (izquierda), la DESC_81 (centro) y la DESC_8 (derecha) sobre la zona de estudio (zona sombreada naranja). ....	13
<b>Figura 5.</b> Composición en RGB (VV, VH, VH/VV) de imágenes adquiridas por el satélite Sentinel-1 de la zona de estudio para la fecha de captura del 08/09/2018 de la pasada DESC_81. ....	14
<b>Figura 6.</b> Fechas de adquisición de imágenes Sentinel-1 (ASC_103, DESC_8 y DESC_81) y Sentinel-2 (S2). ....	15
<b>Figura 7.</b> Composición RGB color natural de las adquisiciones Satélite-2 el día 01/09/2019 mostrando las cuatro cuadrículas que comprenden la zona de estudio. ....	16
<b>Figura 8.</b> Diagrama de flujo de la metodología aplicada en el trabajo. ....	18
<b>Figura 9.</b> Ejemplo de dos reflectores puntuales en la zona de estudio: (a) aerogenerador en el borde de una parcela y (b) edificio. ....	19
<b>Figura 10.</b> Distribución de las parcelas seleccionadas en la zona de estudio. ....	20
<b>Figura 11.</b> Ejemplo de huecos de una y cinco capturas generados en una serie temporal, a partir de la primera, segunda, tercera y última fecha de adquisición. ....	24
<b>Figura 12.</b> Diagrama de flujo del método planteado por Chen et al. (2004). ....	25
<b>Figura 13.</b> Correlación media de parcelas por cultivo entre los índices de Sentinel-1 y Sentinel-2 para las tres pasadas de Sentinel-1. ....	30
<b>Figura 14.</b> Matriz de correlaciones medias de las parcelas del cultivo de cebada entre los índices Sentinel-1 (pasada DESC_81) y Sentinel-2. ....	31
<b>Figura 15.</b> Matrices de correlaciones medias (en valor absoluto) de las parcelas del CE de la cebada en regadío entre los índices de Sentinel-1 y Sentinel-2 en unidades de decibelios (dB, izquierda) y lineales ( $m^2m^{-2}$ , derecha). *Media de la matriz de correlaciones medias. ....	32
<b>Figura 16.</b> Matrices de correlaciones medias (en valor absoluto) de las parcelas del CE de la cebada en regadío entre los índices de Sentinel-1 y Sentinel-2, con datos de toda la serie temporal ( $ST_C$ , izquierda) y calendario fenológico de la cebada ( $ST_{VT}$ , derecha). *Media de la matriz de correlaciones medias. ....	33
<b>Figura 17.</b> Matrices de correlaciones medias de las 75 parcelas del CE de la cebada en regadío entre los índices de Sentinel-1 y Sentinel-2, con filtro SG aplicado para diferente tamaño de ventana: 15 ( $ST_{VT,15}$ , izquierda) y 31 ( $ST_{VT,31}$ , derecha). *Media de la matriz de correlaciones medias. ....	35
<b>Figura 18.</b> Mediana e IQR (sombreado) de las series temporales (01/02/2019 – 31/07/2019) suavizadas mediante el filtro SG (ventana de 31 días) y normalizadas, para los índices de Sentinel-1 (azul) y Sentinel-2 (naranja), para el cultivo de cebada en regadío. Zonas sombreadas: período vegetativo, período reproductivo, llenado y madurez, cosecha. ....	37
<b>Figura 19.</b> Mediana e IQR (sombreado) de las series temporales (01/02/2019 – 31/12/2019) suavizadas mediante el filtro SG (ventana de 31 días) y normalizadas, para los índices de Sentinel-1 (azul) y Sentinel-	

2 (naranja), para el cultivo de trigo en regadío. Zonas sombreadas: período vegetativo, período reproductivo, llenado y madurez, cosecha.....38

**Figura 20.** Mediana e IQR (sombreado) de las series temporales (01/04/2019 – 15/10/2019) suavizadas mediante el filtro SG (ventana de 31 días) y normalizadas, para los índices de Sentinel-1 (azul) y Sentinel-2 (naranja), para el cultivo de girasol. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.....39

**Figura 21.** Mediana e IQR (sombreado) de las series temporales (01/07/2019 – 31/12/2019) suavizadas mediante el filtro SG (ventana de 31 días) y normalizadas, para los índices de Sentinel-1 (azul) y Sentinel-2 (naranja), para el cultivo de maíz. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.....40

**Figura 22.** Mediana e IQR (sombreado) de las series temporales (01/03/2019 – 31/12/2019) suavizadas mediante el filtro SG (ventana de 31 días) y normalizadas, para los índices de Sentinel-1 (azul) y Sentinel-2 (naranja), para el cultivo de la vid. Zonas sombreadas: período vegetativo, período reproductivo, llenado y madurez, cosecha.....41

**Figura 23.** Representación de la mediana e IQR (sombreado) del RMSE para las parcelas del CT de cebada en regadío en función de la duración de los huecos: 1c (un captura sin información), 2c (dos capturas seguidas sin información), etc. hasta 20c, a lo largo de las ventanas temporales definidas. ....42

**Figura 24.** RMSE medio de las parcelas de CT para el cultivo de la cebada en regadío en función del momento en el que se simulan los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.....43

**Figura 25.** RMSE medio de las parcelas de CT para el cultivo del trigo blando en secano (centro) en función del momento en el que se simulan los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.....43

**Figura 26.** RMSE medio de las parcelas de CT para el cultivo del girasol en función del momento en el que se simulan los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.....44

**Figura 27.** RMSE medio de las parcelas de CT para el cultivo del maíz en función del momento en el que se simulan los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.....45

**Figura 28.** RMSE medio de las parcelas de CT para el cultivo de la vid, para una ventana temporal de 01/03/2019 – 31/12/2019, en función del momento en el que se simulan los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.....45

**Figura 29.** Serie temporal NDVI estimada obtenida a partir de datos de índices Sentinel-1 mediante regresión lineal (izquierda) y regresión polinómica de segundo grado (derecha) para el cultivo de la cebada en regadío. \*RMSE Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.....46

**Figura 30.** RMSE medio entre la serie de NDVI original (STi) y las series reconstruidas (STAR<sub>Si</sub>) a partir de índices radar mediante regresiones lineales (RL, línea continua) y polinómicas (RP2, línea discontinua)

para la cebada en regadío, en función de la duración de los huecos. Zonas sombreadas: **germinación y período vegetativo**, **período reproductivo**, **llenado y madurez**, **cosecha**.....48

**Figura 31.** RMSE medio entre la serie de NDVI original (STi) y las series reconstruidas (STAR<sub>S1</sub>) a partir del índice radar RDIV mediante regresiones lineales (RL) y polinómicas (RP2) para la cebada en secano, en función de la duración de los huecos. Zonas sombreadas: **germinación y período vegetativo**, **período reproductivo**, **llenado y madurez**, **cosecha**.....49

**Figura 32.** RMSE medio entre la serie de NDVI original (STi) y las series reconstruidas (STAR<sub>S1</sub>) a partir del índice radar RDIV mediante regresiones lineales (RL) y polinómicas (RP2) para el trigo en regadío (arriba) y en secano (abajo), en función de la duración de los huecos. Zonas sombreadas: **germinación y período vegetativo**, **período reproductivo**, **llenado y madurez**, **cosecha**.....50

**Figura 33.** RMSE medio entre la serie de NDVI original (STi) y las series reconstruidas (STAR<sub>S1</sub>) a partir los índices radar RDIF y RDIV mediante regresiones lineales (RL) y polinómicas (RP2) para el girasol, en función de la duración de los huecos. Zonas sombreadas: **germinación y período vegetativo**, **período reproductivo**, **llenado y madurez**, **cosecha**.....51

**Figura 34.** RMSE medio entre la serie de NDVI original (STi) y las series reconstruidas (STAR<sub>S1</sub>) a partir el índice radar VH mediante regresiones lineales (RL) y polinómicas (RP2) para el maíz, en función de la duración de los huecos. Zonas sombreadas: **germinación y período vegetativo**, **período reproductivo**, **llenado y madurez**, **cosecha**.....51

**Figura 35.** RMSE medio entre la serie de NDVI original (STi) y las series reconstruidas (STAR<sub>S1</sub>) a partir del índice radar RDIF mediante regresiones lineales (RL) y polinómicas (RP2) para la vid, en función de la duración de los huecos. Zonas sombreadas: **germinación y período vegetativo**, **período reproductivo**, **llenado y madurez**, **cosecha**.....52

**Figura 36.** RMSE medio entre la serie de NDVI original (STi) y las series reconstruidas a partir de la propia serie de Sentinel-2 (STAR<sub>S2</sub>) para la cebada en regadío (arriba) y secano (abajo), en función de la duración de los huecos. Zonas sombreadas: **germinación y período vegetativo**, **período reproductivo**, **llenado y madurez**, **cosecha**.....53

**Figura 37.** RMSE medio entre la serie de NDVI original (STi) y las series reconstruidas a partir de la propia serie de Sentinel-2 (STAR<sub>S2</sub>) para el trigo en regadío (arriba) y secano (abajo), en función de la duración de los huecos. Zonas sombreadas: **germinación y período vegetativo**, **período reproductivo**, **llenado y madurez**, **cosecha**.....54

**Figura 38.** RMSE medio entre la serie de NDVI original (STi) y las series reconstruidas a partir de la propia serie de Sentinel-2 (STAR<sub>S2</sub>) para el girasol, en función de la duración de los huecos. Zonas sombreadas: **germinación y período vegetativo**, **período reproductivo**, **llenado y madurez**, **cosecha**.....55

**Figura 39.** RMSE medio entre la serie de NDVI original (STi) y las series reconstruidas a partir de la propia serie de Sentinel-2 (STAR<sub>S2</sub>) para el maíz, en función de la duración de los huecos: 1c- 5c (izquierda) y 5c – 20c (derecha), en función de la duración de los huecos. Zonas sombreadas: **germinación y período vegetativo**, **período reproductivo**, **llenado y madurez**, **cosecha**.....55

**Figura 40.** RMSE medio entre la serie de NDVI original (STi) y las series reconstruidas a partir de la propia serie de Sentinel-2 (STAR<sub>S2</sub>) para el maíz, en función de la duración de los huecos. Zonas sombreadas: **germinación y período vegetativo**, **período reproductivo**, **llenado y madurez**, **cosecha**.....56

**Figura 41.** Resumen de las técnicas que presentan mejores resultados (menor RMSE) en función de la fecha en la que ocurre el hueco y la duración de este, correspondiendo el color azul a Sentinel-2 y el color verde a Sentinel-1 (regresión lineal, RDIV), para el cultivo de cebada en secano. Fechas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha. ....57

**Figura 42.** Resumen de las técnicas que presentan mejores resultados (menor RMSE) en función de la fecha en la que ocurre el hueco y la duración de este, correspondiendo el color azul a Sentinel-2 y el color verde a Sentinel-1 (regresión lineal, RDIV), para el cultivo de cebada en regano. Fechas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha. ....57

**Figura 43.** Resumen de las técnicas que presentan mejores resultados (menor RMSE) en función de la fecha en la que ocurre el hueco y la duración de este, correspondiendo el color azul a Sentinel-2 y el color verde a Sentinel-1 (regresión lineal, RDIV), para el cultivo de girasol. Fechas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha. ....57

**Figura 44.** Resumen de las técnicas que presentan mejores resultados (menor RMSE) en función de la fecha en la que ocurre el hueco y la duración de este, correspondiendo el color azul a Sentinel-2 y el color verde a Sentinel-1 (regresión lineal, VH), para el cultivo de maíz. Fechas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha. ....58

**Figura 45.** Resumen de las técnicas que presentan mejores resultados (menor RMSE) en función de la fecha en la que ocurre el hueco y la duración de este, correspondiendo el color azul a Sentinel-2 y el color verde a Sentinel-1 (regresión lineal, RDIV en regadío, regresión lineal VH en secano), para el cultivo de trigo. Fechas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha. ....58

**Figura 46.** Resumen de las técnicas que presentan mejores resultados (menor RMSE) en función de la fecha en la que ocurre el hueco y la duración de este, correspondiendo el color azul a Sentinel-2 y el color verde a Sentinel-1 (regresión polinómica, RDIF), para el cultivo de la vid. Fechas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha. ....59

## ÍNDICE DE TABLAS

<b>Tabla 1.</b> Características de la misión Sentinel-1.....	2
<b>Tabla 2.</b> Características de la misión Sentinel-2.....	4
<b>Tabla 3.</b> Calendario de cultivos principales de la zona de estudio divididos en: cultivos de invierno (cebada y trigo blando), de verano (girasol y maíz) y permanente (viñedo vinificación – Vid- ). .....	12
<b>Tabla 4.</b> Especificaciones de las bandas espectrales de Sentinel-2.....	15
<b>Tabla 5.</b> Número de parcelas y superficie ocupada por los cultivos de interés.....	17
<b>Tabla 6.</b> Definiciones y fórmulas de los índices de vegetación ópticos calculados. ....	20
<b>Tabla 7.</b> Definiciones y fórmulas de los índices radar calculados. ....	21
<b>Tabla 8.</b> Ventanas temporales aplicadas en la serie temporal de estudio para cada cultivo. ....	23
<b>Tabla 9.</b> Correlación media entre la opción de cálculo a nivel de pixel y a nivel de parcela, para los 10 índices ópticos propuestos. ....	28
<b>Tabla 10.</b> NRMSE medio entre la opción de cálculo a nivel de pixel y a nivel de parcela, para los 10 índices ópticos propuestos. ....	28
<b>Tabla 11.</b> Correlación media entre la opción de cálculo a nivel de pixel y a nivel de parcela, para los 7 índices radar propuestos .....	29
<b>Tabla 12.</b> NRMSE medio entre la opción de cálculo a nivel de pixel y a nivel de parcela, para los 7 índices radar propuestos. ....	29
<b>Tabla 13.</b> Grupos de índices de Sentinel-1 con comportamiento de correlación similar frente a índices de Sentinel-2. ....	30
<b>Tabla 14.</b> Grupos de índices de Sentinel-2 con comportamiento de correlación similar frente a índices de Sentinel-1. ....	31
<b>Tabla 15.</b> Datos de interés de las matrices de correlaciones medias de comparación de unidades (dB vs. $m^2m^{-2}$ ) para cada cultivo, indicando las unidades con mayor media de correlaciones y la mejor combinación de índices.....	33
<b>Tabla 16.</b> Datos de interés de las matrices de correlaciones medias de comparación de ventanas temporales ( $ST_C$ vs. $ST_{VT}$ ) para cada cultivo, indicando la mejor opción y la mejor combinación de índices. ....	34
<b>Tabla 17.</b> Datos de interés de las matrices de correlaciones medias de comparación de tamaño de ventanas aplicadas mediante filtro SG: sin filtro, 15 y 31 días ( $ST_{VT}$ vs. $ST_{VT,15}$ vs. $ST_{VT,31}$ ) para cada cultivo, indicando la mejor opción y la mejor combinación de índices.....	35
<b>Tabla 18.</b> Resumen de las medianas de RMSE en NDVI para las parcelas de CT de los cultivos de interés, en la ventana temporal definida para cada cultivo.....	42
<b>Tabla 19.</b> Resumen de valores de RMSE entre la serie temporal original de NDVI y la serie de referencia de NDVI obtenida a partir de los diferentes índices radar mediante regresión lineal (RL) y regresión polinómica de 2º grado (RP2) para los cultivos de interés. ....	47



## LISTADO DE ABREVIATURAS

<i>ASC</i>	Ascendente
<i>CE</i>	Conjunto de entrenamiento
<i>CI-RE</i>	Chlorophyll index Red-Edge CIRE
<i>CT</i>	Conjunto de test
<i>dB</i>	Decibelios
<i>DESC</i>	Descendente
<i>IDPDD</i>	Inverse Dual-Pol Diagonal Distance
$m^2m^{-2}$	unidades lineales
<i>MCARI</i>	Modified Chlorophyll Absorption in Ratio Index
<i>MTCI</i>	Meris Terrestrial Chlorophyll Index
<i>NDMI</i>	Normalized Difference Vegetation Index
<i>NDVI</i>	Normalised Difference Vegetation Index
<i>NDVire</i>	Red-edge NDVI
<i>NIR</i>	Infrarrojo cercano
<i>NRMSE</i>	Raíz del error cuadrático medio normalizado
<i>RDVI</i>	Renormalized Difference Vegetation Index
<i>RMSE</i>	raíz del error cuadrático medio
<i>RVIVV</i>	Radar Vegetation Index
<i>S1</i>	Sentinel-1
<i>S2</i>	Sentinel-2
<i>SAR</i>	Radar de Apertura Sintética
<i>SAVI</i>	Soil Adjusted Vegetation Index
<i>SeLI</i>	Sentinel-2 LAI index
<i>SG</i>	Savitzky-Golay
<i>SLAVI</i>	Specific Leaf Area Vegetation Index
<i>ST</i>	Series temporales originales
<i>STA</i>	Series temporales artificiales (STi5 con huecos)
<i>STAi</i>	Serie temporal artificial interpolada
<i>STAR<sub>S1</sub></i>	Serie temporal artificial rellena a partir de datos de S1
<i>STAR<sub>S2</sub></i>	Serie temporal artificial rellena a partir de datos de S2
<i>ST<sub>C</sub></i>	Serie temporal completa
<i>STi</i>	Series temporales originales interpoladas a paso diario
<i>STi5</i>	Series temporales originales interpoladas a paso diario, con paso de 5 días
<i>ST<sub>VT</sub></i>	Serie temporal con ventana temporal aplicada
<i>ST<sub>VTSG</sub></i>	Serie temporal suavizada mediante el filtro de Savitzky-Golay
<i>SWIR</i>	Infrarrojo medio de onda corta
<i>VDDPI</i>	Vertical Dual depolarization Index
<i>VI</i>	Índice de vegetación
<i>VT</i>	Ventana temporal



## 1. INTRODUCCIÓN

La teledetección es la técnica mediante la cual se obtiene información de cubiertas y elementos presentes en la superficie terrestre a través del análisis de imágenes adquiridas por un sensor alojado en una plataforma espacial/aérea (Cambell & Wynne, 2011). Los sensores mediante los cuales se obtiene la información suelen ser clasificados en dos grandes grupos: pasivos y activos. Los primeros captan la energía electromagnética procedente de las cubiertas terrestres (bien reflejada de la radiación solar o emitida por la temperatura propia de la superficie), mientras que los sensores activos emiten un haz energético hacia la superficie terrestre que posteriormente recogen, tras la reflexión de este sobre la superficie observada (Chuvieco, 1995). Desde el punto de vista de la agricultura, los sensores más interesantes son, entre los pasivos, los ópticos multispectrales y, entre los activos, los SAR (radar de apertura sintética) (ver apartado 1.2).

### 1.1. Programa Copernicus

Copernicus es el Programa de Observación de la Tierra de la Unión Europea que ofrece servicios de información basados en datos de observación de esta obtenidos mediante satélites y en datos *in situ* (no espaciales), coordinado y gestionado por la Comisión Europea y ejecutado por los Estados miembros y la Agencia Espacial Europea (ESA), entre otros (CE, 2020). El principal objetivo del programa es proporcionar información que ayude a los proveedores de servicios, las administraciones públicas y otras organizaciones internacionales a mejorar la calidad de vida de la ciudadanía europea, prestando servicios de información de acceso gratuito y abierto y facilitando datos globales en tiempo cuasirreal.

Para ello, Copernicus cuenta con un conjunto de satélites propios (los Sentinel) complementados con misiones participantes e información procedente de sistemas *in situ* como las estaciones en tierra. Entre la familia de satélites de Sentinel se encuentran los satélites Sentinel-1 y Sentinel-2, cuyos productos derivados (disponibles para los usuarios de manera gratuita, íntegra y abierta) serán los que se utilicen en el presente trabajo.

La misión Sentinel-1 comprende una constelación de dos satélites gemelos (Sentinel-1A y Sentinel-1B, lanzados el 3 de abril de 2014 y el 25 de abril de 2016, respectivamente) que capturan imágenes SAR en la banda C con diferentes resoluciones (hasta 5 m) y coberturas (hasta 400 km), permitiendo adquirir datos durante el día o la noche, en casi cualquier condición climática, debido a las ventajas que supone trabajar con sensores activos que no requieren la energía solar y que emiten en longitudes de onda que no se ven obstaculizadas por la nubosidad o la falta de iluminación. Cada uno de los satélites gemelos de la misión está situado en una órbita casi polar heliosíncrona, con un ciclo de repetición de 12 días. Al compartir el mismo plano orbital, con una diferencia de base de 180°, con ambos satélites en funcionamiento, el periodo de revisita se reduce

a 6 días. Sentinel-1 es un sistema SAR de polarización dual que preserva la fase, es decir, puede transmitir una señal en horizontal (H) o vertical (V) y recibirla en polarización H o V.

Entre los múltiples servicios que presenta la misión, cabe destacar que Sentinel-1 permite obtener información sobre los océanos y el Ártico gracias a la capacidad de la misión de proporcionar observaciones en casi cualquier condición, así como beneficiar a los usuarios interesados en una gran variedad de aplicaciones como la agricultura, selvicultura y coberturas el suelo gracias a sus productos de polarización doble.

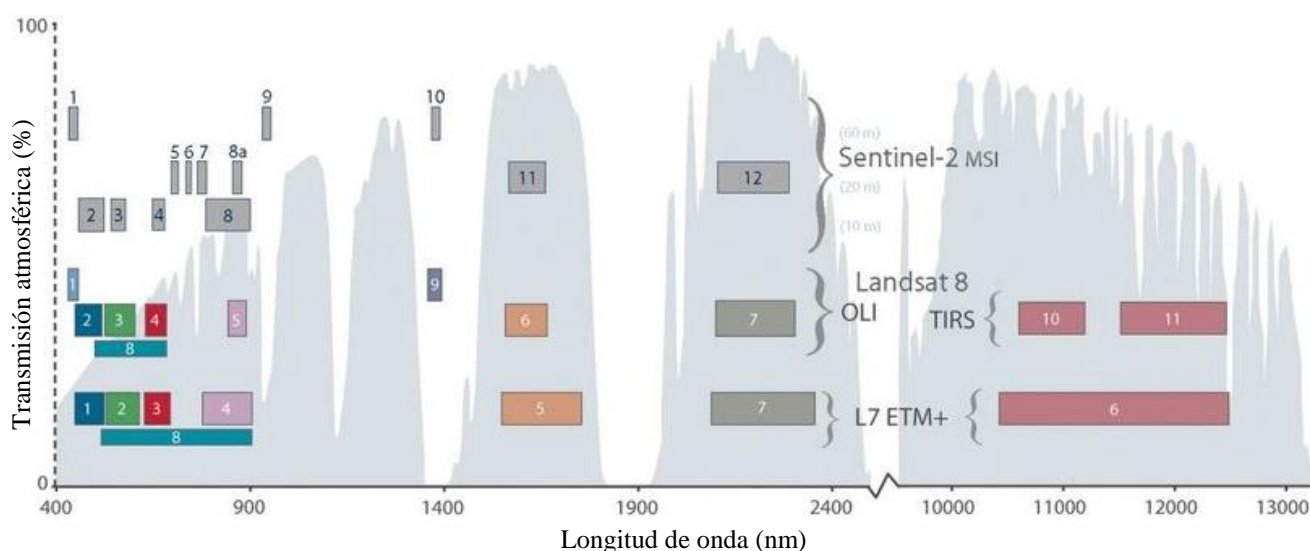
Las características más importantes de la misión Sentinel-1 quedan resumidas en la **Tabla 1**.

**Tabla 1.** Características de la misión Sentinel-1.

<b>Fecha de lanzamiento</b>	Sentinel-1A: 3 de abril de 2014 Sentinel-1B: 25 de abril de 2016
<b>Objetivos de la misión</b>	<ul style="list-style-type: none"> <li>• Monitoreo terrestre de bosques, agua, suelo y agricultura</li> <li>• Apoyo cartográfico de emergencia en caso de desastres naturales</li> <li>• Vigilancia marina del medio marítimo</li> <li>• Observaciones del hielo marino y seguimiento de los icebergs</li> <li>• Producción de cartas de hielo de alta resolución</li> <li>• Pronóstico de las condiciones del hielo en el mar</li> <li>• Mapeo de derrames de petróleo</li> <li>• Detección de embarcaciones marinas</li> <li>• Seguimiento del cambio climático.</li> </ul>
<b>Órbita</b>	<ul style="list-style-type: none"> <li>• Órbita circular heliosíncrona, casi polar</li> <li>• 693 km de altura</li> <li>• Inclinación 98,18 °</li> <li>• Periodo revisita: 12 días en el Ecuador con un satélite (175 órbitas/ciclo)</li> </ul>
<b>Radar</b>	<ul style="list-style-type: none"> <li>• Radar de apertura sintética (SAR) en banda-C.</li> <li>• Frecuencia: 5,405 GHz.</li> <li>• Polarizaciones: VV+VH, HH+HV, HH, VV</li> <li>• Ángulos de incidencia: 20°-45°</li> </ul>
<b>Modos de trabajo</b>	<ul style="list-style-type: none"> <li>• Strip Map (SM): 80 km ancho, 5 x 5 m resolución espacial.</li> <li>• Interferometric Wide Swath (IW): 250 km ancho, 5 x 20 m resolución espacial.</li> <li>• Extra Wide Swath (EW): 400 km ancho, 20 x 40 m resolución espacial.</li> <li>• Wave-Mode (WM): viñetas de 20 x 20 km, 5 x 5 m resolución espacial.</li> </ul>
<b>Productos derivados</b>	<ul style="list-style-type: none"> <li>• Nivel 0 (Raw): datos en bruto.</li> <li>• Nivel 1.1 (Single Look Complex SLC): imágenes en geometría Slant Range, con el espaciado original en rango y acimut.</li> <li>• Nivel 1.5 (Ground Range Detected GRD): imágenes transformadas a Ground Range y con procesado multilook.</li> <li>• Nivel 2.0 (Ocean): productos geofísicos georreferenciados.</li> </ul>

La misión Sentinel-2 (S2) también está compuesta por dos satélites gemelos (Sentinel-2A y Sentinel-2B lanzados el 23 de junio de 2015 y 7 de marzo de 2017, respectivamente) que facilitan imágenes ópticas de alta resolución, proporcionando imágenes para el seguimiento de la vegetación, el suelo y las masas de agua, vías de navegación, etc., así como para los servicios de emergencia. Del mismo modo que la misión Sentinel-1, cada uno de los satélites gemelos de la misión S2 está situado en una órbita casi polar heliosíncrona, con un ciclo de repetición de 10 días (con ambos satélites en funcionamiento, el periodo de revisita se reduce a 5 días), capturando

imágenes multiespectrales, en 13 bandas (**Figura 1**), con diferentes resoluciones espaciales (cuatro bandas a 10 m, seis bandas a 20 m y tres bandas a 60 m) y un ancho de barrido de 290 km. A diferencia de la misión S1, las longitudes de onda en las que trabaja S2 se ven obstaculizadas por la nubosidad, además de ser un sensor óptico que requiere de la energía solar, por lo que se verá también obstaculizado por la falta de iluminación. Frente a este problema, Sentinel-2 introdujo varias mejoras respecto a otros satélites como Landsat 7 y 8, reduciendo el ancho de varias bandas espectrales (las bandas del *red edge* a una resolución espacial de 20 m y NIR), limitando así la influencia de los componentes atmosféricos, incluido el valor de agua. Las nuevas bandas introducidas (B5, B6, B7 y B8A) fueron diseñadas para poder representar la curva del *red-edge* evitando la contaminación por vapor de agua. Además, también incorpora las bandas B1 para el contenido de aerosoles, B9 para la detección de nubes cirrus y B10 para la detección de vapor de agua, con el fin de poder adquirir información del estado de la atmósfera y así poder realizar correcciones atmosféricas del resto de las bandas, estimando los valores de reflectancia a nivel de la superficie terrestre.



**Figura 1.** Comparación de las bandas espectrales de Sentinel-2 y Landsat 8 y 9. Fuente: Poursanidis & Chrysoulakis (2017)

Las características más importantes de la misión Sentinel-2 quedan resumidas en la **Tabla 2**.

**Tabla 2.** Características de la misión Sentinel-2.

<b>Fecha de lanzamiento</b>	Sentinel-2A: 23 de junio de 2015 Sentinel-2B: 7 de marzo de 2017
<b>Áreas temáticas y servicios</b>	<ul style="list-style-type: none"> <li>• Monitoreo terrestre</li> <li>• Gestión de emergencias</li> <li>• Seguridad</li> <li>• Cambio climático</li> </ul>
<b>Objetivos de la misión</b>	<ul style="list-style-type: none"> <li>• Adquisiciones globales sistemáticas de imágenes multiespectrales de alta resolución.</li> <li>• Continuidad en la adquisición de imágenes multiespectrales.</li> <li>• Proporcionar datos de observación para la próxima generación de productos operativos (mapas de cobertura terrestre, cambios de tierra, etc.)</li> </ul>
<b>Órbita</b>	<ul style="list-style-type: none"> <li>• Órbita circular heliosíncrona, casi polar</li> <li>• 786 km de altura</li> <li>• Inclinación 98,62 °</li> <li>• Periodo revisita: 10 días en el Ecuador con un satélite</li> </ul>
<b>Productos derivados</b>	<ul style="list-style-type: none"> <li>• Nivel-1C: reflectancias en el techo de la atmósfera en geometría cartográfica.</li> <li>• Nivel-2A: reflectancias en la superficie en geometría cartográfica.</li> </ul>

## 1.2. Teledetección y agricultura

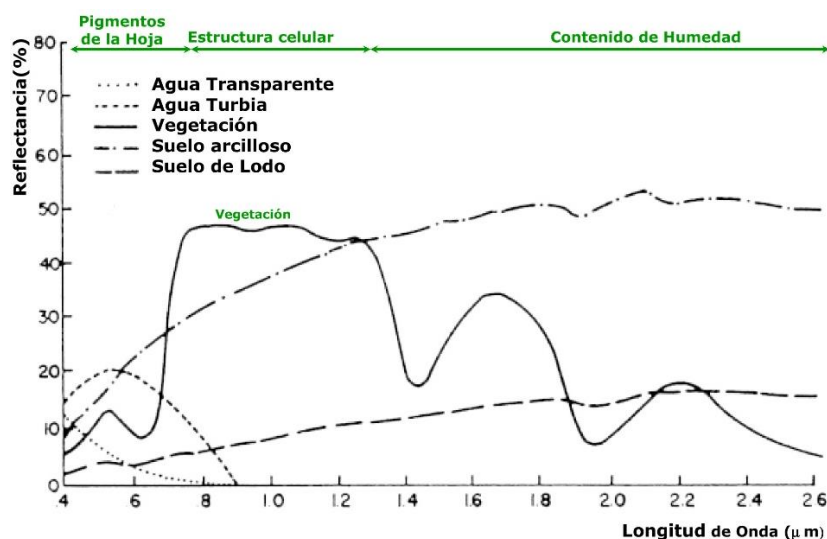
Como se ha ido comentando, la evolución de la tecnología de la teledetección ha desarrollado nuevos sensores de alto rendimiento con resoluciones espaciales, espectrales y temporales más altas. Aplicar la tecnología de la teledetección es, por lo tanto, un componente clave en la agricultura ya que permite generar imágenes y datos espectrales de gran volumen que pueden servir como apoyo la toma de decisiones en la gestión agrícola, como en la fertilización, el riego o el manejo de plagas (Huang et al., 2018; Liaghat & Balasundram, 2010). Los productos derivados de los sensores ópticos utilizados en la agricultura, consisten normalmente en reflectancias del visible, NIR y regiones del SWIR, que pueden ser interpretadas en términos de parámetros físicos (contenido de pigmentos, estructura celular de las hojas, contenido de humedad, etc.) a través de diferentes índices de vegetación, así mismos los productos derivados de los sensores radar, permiten extraer parámetros estructurales de la vegetación (Mulla, 2013).

### 1.2.1. Índices ópticos y radar

Los índices de vegetación (VI, del inglés *Vegetation Indices*) consisten en la transformación aritmética de dos o más bandas, diseñados para reducir la dimensionalidad de la información y contribuir al conocimiento de las propiedades vegetales. Por otro lado, los VI permiten realizar comparaciones multitemporales y espaciales fidedignas de atributos vegetales como la actividad fotosintética terrestre y de las variaciones estructurales de las cubiertas (Huete et al., 2002). El concepto de VI nació en la década de los 70, siendo, desde entonces, una herramienta muy utilizada tanto para la evaluación del estado de la vegetación (el índice de área foliar, la cobertura o la fenología, entre otros), así como para la evapotranspiración o la productividad de un cultivo (Glenn et al., 2008). Al ser simples transformaciones aritméticas de las bandas espectrales, los VIs son calculados directamente sin ningún sesgo ni suposición con respecto a la clasificación de

las coberturas terrestres, el tipo de suelo o las condiciones climáticas, permitiendo monitorizar variaciones estacionales e interanuales de largo plazo de los parámetros comentados anteriormente (Frampton et al., 2013; Glenn et al., 2008; Huete et al., 2002).

Los índices de vegetación derivados de los valores de reflectancia adquiridos por sensores ópticos son ampliamente utilizados en el estudio de las variaciones del estado fenológico y las propiedades biofísicas de la vegetación (Sonobe et al., 2018). Los primeros VIs contrastaban la alta reflectividad en la banda del infrarrojo cercano (NIR, del inglés *near-infrared*, correspondiente a la banda 8 de Sentinel-2, **Figura 2**), debido a la estructura celular del mesófilo de las hojas, con la banda del rojo (correspondiente a la banda 4 de Sentinel-2), donde hay una gran absorción de la radiación del espectro electromagnético debido al contenido de clorofila de esta, con el fin de cuantificar el *verdor* o vigor de la vegetación (Frampton et al., 2013).



**Figura 2.** Curva de respuesta espectral para distintas coberturas. *Fuente:* (Chuvieco, 2006)

Este ratio (posteriormente llamado SR, del inglés *Simple Ratio*), es la base del índice de vegetación NDVI (del inglés *Normalized Difference Vegetation Index*), considerado uno de los VIs más comúnmente utilizado, propuesto por primera vez por Rouse et al. (1973). El NDVI es un ratio normalizado de las bandas del NIR y del rojo. De acuerdo con Glenn et al. (2008), este índice se fundamenta en el siguiente principio: la clorofila a y b, en hojas verdes, absorbe de manera importante la radiación electromagnética en la banda del rojo, con su máxima absorción en la longitud de onda de 690 nm, mientras que las paredes celulares reflejan gran parte de la radiación en la región del NIR, con el máximo cerca de los 850 nm. El resultado, por lo tanto, muestra un fuerte contraste de la absorción entre las dos bandas, dando valores normalizando entre -1 y 1, correspondiendo valores altos positivos a vegetación densa, valores bajos, también positivos, a suelos, y valores negativos a masas de agua (debido a su alta absorción en el infrarrojo cercano). Sobre este fundamento también se basan otros VIs que se aplicarán en este trabajo, como el índice SAVI (*Soil Adjusted Vegetation Index*), propuesto por primera vez por Huete



(1988), sobre el que introduce un factor de ajuste de suelo, que tiene como objetivo minimizar las influencias del suelo subyacente sobre la reflectancia de las cubiertas vegetales, con un factor que oscila entre 0 (zonas con gran densidad vegetal) y 1 (baja densidad vegetal) (Huete, 1988; Qi et al., 1994). El SAVI es interesante principalmente en cubiertas vegetales que dejan libre mucho espacio en el suelo (e.g. cultivos como viñas u olivos). Otro índice de vegetación que se plantea, teniendo en cuenta únicamente las bandas del NIR y del rojo, es el RDVI (*Renormalized Difference Vegetation Index*) (Roujean & Breon, 1995), con el que se pretenden corregir las saturaciones del índice NDVI frente a cubiertas vegetales densas y compuestas por múltiples capas (Haboudane et al., 2004; Main et al., 2011).

Se plantean también diferentes VIs que incorporan la banda del infrarrojo medio de onda corta (SWIR, del inglés *short-wave infrared*, correspondiente a la banda 11 de Sentinel-2) a las dos bandas mencionadas en los índices anteriores. Un ejemplo de ello, propuesto Cibula et al. (1992), es el índice NDMI (Normalized Difference Moisture Index), que muestra la capacidad de absorción de agua de la vegetación. De acuerdo con Wilson & Sader (2002), la diferencia entre el NIR y el SWIR se ve reflejada en la capacidad de la banda de este último para absorber el agua, ya que, en una vegetación vigorosa, la banda del NIR tiene una alta reflectancia comparada con valores más reducidos en la banda del SWIR, debidos a la absorción del agua. Otro VI que relaciona estas tres bandas, propuesto por Lymburner et al. (2000), es el índice SLAVI (Specific Leaf Area Vegetation Index). Este índice surge de la necesidad de crear relaciones entre la variable superficie foliar específica (SLA, del inglés *Specific Leaf Area*) y los VIs, obteniendo valores de SLAVI bajos cuando la reflectancia en el NIR sea baja (asociada a valores bajos de SLA). De la misma manera, valores altos tanto en la banda del SWIR como del rojo están asociados a valores de SLA bajos, por lo que cuanto mayores sean estos, menor será el valor de SLAVI (Lymburner et al., 2000; Zhang & Ke, 2016).

Los últimos índices que se plantean para el espectro óptico se denominan índices de banda estrecha, ya que están relacionados con las bandas del *red edge* (RE, correspondientes a las bandas 5, 6 y 7 de Sentinel-2) del espectro electromagnético, combinado con la banda del NIR o la del rojo. Por un lado, el índice de clorofila – Red Edge (CI-RE), propuesto por Gitelson et al. (2003). Como su nombre indica, es un VI que muestra el contenido de clorofila de la vegetación, ya que la profundidad de penetración de la radiación en una hoja debido a su contenido en clorofila en la banda del RE es bastante superior a las de la banda del azul y el rojo (Gitelson, Gritz, et al., 2003; Gitelson, Vina, et al., 2003). Por otro lado, el índice MCARI (*Modified Chlorophyll Absorption in Ratio Index*) (Daughtry et al., 2000), indica la profundidad de la absorción de la clorofila a 670 nm relativa a la reflectancia a 700 nm y 550 nm (Daughtry et al., 2000; Haboudane et al., 2004). Parecido al MCARI, Dash and Curran (2004) plantearon el índice MTCI (*MERIS Terrestrial Chlorophyll Index*), ideado en un principio para imágenes MERIS, ha acabado siendo un índice



utilizado debido a su buena correlación con el contenido en nitrógeno en cultivos agrícolas. Pasqualoto et al (2019) planteó SeLI (Sentinel-2 LAI Index) como un índice normalizado que utiliza la B5 de Sentinel-2 para equilibrar la influencia de la fuerte absorción de la clorofila. Finalmente se plantea el índice NDVI<sub>re</sub> (Red-edge NDVI) (Gitelson & Merzlyak, 1994), muy empleado en la agricultura de precisión, basado en el mismo principio que el índice NDVI, pero centrándose en el inicio y en el final de la curva de máxima pendiente de la signatura espectral de la vegetación.

Tanto los índices que se acaban de comentar como la información adquirida mediante teledetección óptica puede ser combinada y complementada con datos obtenidos por medio de teledetección radar (y los índices derivados de esta). Como ya se ha mencionado con anterioridad, una de las principales ventajas del radar es que, a diferencia de los sensores ópticos, que ofrecen información acerca de la reflectividad de las cubiertas (debidas principalmente a su composición), los sensores radar son sensibles a otras características de la vegetación, como la geometría (forma, tamaño y orientación) o las características dieléctricas de las cubiertas, permitiendo aportar información complementaria en situaciones en las que los sensores ópticos resulten insensibles a ciertas características de interés de las cubiertas. Por ello se plantea un primer VI, el RVI (del inglés *Radar Vegetation Index*), propuesto por Jordan (1969). De acuerdo con Yunjin Kim & Van Zyl (2009), el RVI varía entre el 0 y el 1, siendo los valores cercanos a 0 señal de dispersión superficial y los cercanos a 1 de retrodispersión de volumen, y por lo tanto, indicadores del desarrollo de la cubierta vegetal. Sin embargo, puede observarse que este índice tiene como limitación la necesidad de trabajar con polarización cuádruple (HH, VV, HV, VH), por lo que se plantean índices RVI<sub>HH</sub> (Trudel et al., 2012) y RVI<sub>VV</sub> (Nasirzadehdizaji et al., 2019) adaptados a configuraciones de polarización doble más habituales a día de hoy (e.g. Sentinel-1).

Existe también un índice con el cual se obtiene el grado de depolarización del pulso vertical incidente de un objeto observado, denominado VDDPI (del inglés *Vertical dual de-polarization index*) propuesto por Periasamy (2018). En el mismo artículo se proponen otros índices, como el IDPDD (del inglés *Inverse dual-Pol Diagonal Distance*), basado en la representación del gráfico de retrodispersión entre  $\sigma_{VV}^0$  y  $\sigma_{VH}^0$ , donde se analiza la diagonal generada entre el punto de  $\sigma_{VV(max)}^0$  y  $\sigma_{VH(max)}^0$ , denominada *Límite suelo-vegetación (a-b)*.

### 1.3. Técnicas de relleno de huecos

Como se ha ido comentando, la continuidad de la disponibilidad de productos ópticos puede verse interrumpida por la presencia de nubes, así como por fallos en el sistema (Chen et al., 2006), reduciendo la utilidad de estas fuentes de datos para modelar y monitorizar la superficie terrestre (Kandasamy et al., 2014). Las técnicas de relleno de huecos hacen uso de las observaciones válidas para estimar la falta de datos, suponiendo que la trayectoria estacional del crecimiento de

la vegetación puede ajustarse mediante un modelo matemático (Liu et al., 2017). Hasta el momento se han propuesto una amplia variedad de enfoques para realizar dicho relleno, clasificados de acuerdo con Liu et al. (2017). en cuatro tipos diferentes, considerando técnicas que tengan en cuenta la trayectoria del crecimiento de la vegetación, la resistencia al ruido y la estabilidad de la curva:

- Basadas en umbrales: técnicas que establecen un umbral para controlar la suavidad de la serie reconstruida, como la técnica del mejor algoritmo de extracción de índice (BISE, por Vivoy et al. (1992), el BISE modificado (Lovell & Graetz, 2001), la iteración de valor medio (MVI, por Ma & Veroustraete (2006)) o la interpolación iterativa para la reconstrucción de datos (IDR, por Julien & Sobrino (2010)).
- Basados en filtros: técnicas que definen un filtro para llenar los vacíos con las observaciones válidas en una ventana móvil, como el filtro Savitzky-Golay (SG, Chen et al. (2004)), el filtro de peso variable (Zhu et al., 2012) o el filtro de regresión local adaptado (Moreno et al., 2014).
- Basados en ajuste de curvas: técnicas que ajustan la trayectoria estacional del crecimiento de la vegetación con una curva matemática, como la función logística doble (DL, por Beck et al., (2006)), la función gaussiana asimétrica (Jönsson & Eklundh, 2002), técnica basada en Fourier (*Fourier*, por Dash et al. (2010)) o ajuste local limitado de un *spline* cúbica (LACC, por J. M. Chen et al. (2006)).
- Otras técnicas como la transformada de Wavelet (Lu et al., 2007), la asimilación de datos (Gu et al., 2009) o el suavizado de Whittaker (*Whit*, por Atzberger & Eilers (2011)).

En ese mismo estudio (Liu et al., 2017), se analizan seis de las técnicas mencionada globalmente, entre ellos el filtro SG propuesto por Chen et al. (2004), píxel por píxel, basadas en un NDVI de referencia generado a partir de observaciones en una serie temporal de 2000-2014. Se evaluaron las técnicas de manera integral, incluyendo el efecto de relleno en huecos simulados (generados de forma artificial), así como en series de huecos en observaciones reales, concluyendo que es una técnica que presenta buenos resultados si las nubes, la nieve, así como otros problemas de comportamiento son previamente detectados.

Otra opción para el relleno de huecos es a partir de información externa, como la proporcionada por Sentinel-1, planteada en diferentes estudios (Caballero et al., 2018; Campos-Taberner et al., 2019; Orynbaikyzy et al., 2020; Pipia et al., 2019; Steinhausen et al., 2018; Van Tricht et al., 2018; Veloso et al., 2017), en los que se analiza las ventajas que supondría la fusión de datos ópticos y radar con diferentes fines (clasificación de cultivos, estimación de LAI, analizar comportamientos temporales de la vegetación y relleno de huecos, entre otros).

Joshi et al. (2016) analiza la utilidad de integrar datos de teledetección y radar a través de la revisión de 112 estudios. De los estudios revisados que evaluaban la ventaja de fusión de datos, la gran mayoría reveló (28 de 32) que la fusión de datos proporciona resultados con mayor precisión que el uso de cualquiera de los conjuntos individualmente.

Aunque se haya visto que es un campo con mucha actividad investigadora, se ha considerado que resulta de interés aplicar y comparar algunas de las técnicas, en este caso en particular, en cultivos agrícolas que sean representativos de la zona de estudio.

## 2. OBJETIVOS

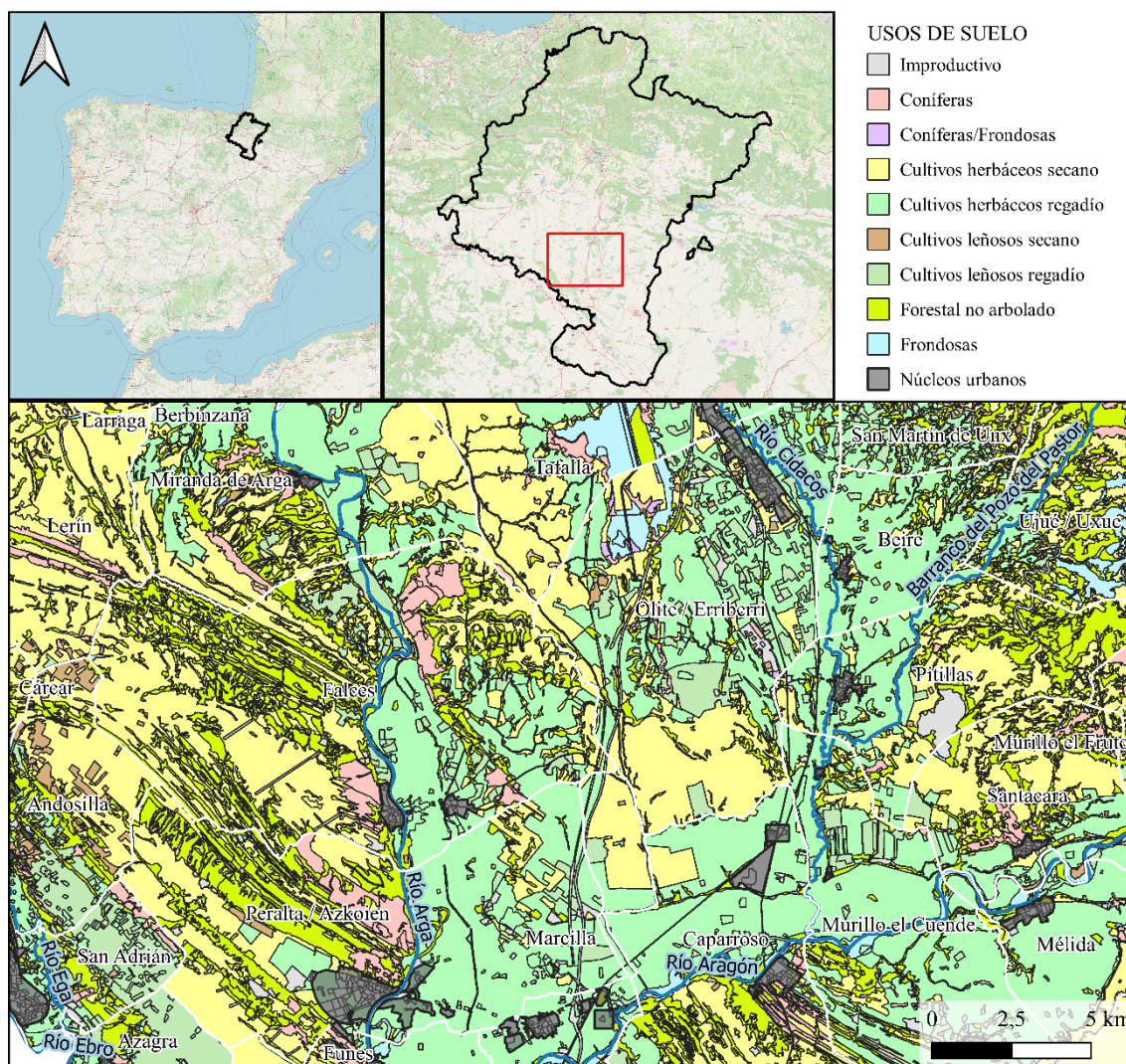
El objetivo general del presente trabajo consiste en evaluar técnicas de relleno de huecos en series temporales de índices de vegetación obtenidos por Sentinel-2. Para ello, se plantean técnicas de relleno con (a) datos procedentes de Sentinel-1 y con (b) datos de la propia serie temporal de Sentinel-2, poniendo el foco en diferentes cultivos de interés. Se analizan diferentes métodos de relleno y se evalúa su implementación en función del momento y duración de los huecos generados artificialmente en una serie temporal. Se definen, además, una serie de objetivos específicos:

- Analizar las diferencias en el cálculo de índices de vegetación a nivel de pixel y a nivel de objeto (parcela) en las series temporales de Sentinel-1 y Sentinel-2.
- Calcular índices de vegetación de Sentinel-1 y Sentinel-2 que permitan extraer información acerca del cultivo, analizando también la correlación entre estos, agrupando aquellos que aporten información similar.
- Interpretar y analizar el comportamiento reflexivo y retrodispersivo de los cultivos principales de la zona de estudio.
- Implementar y comparar distintas técnicas de relleno de huecos y seleccionar las idóneas para cada caso.

### 3. MATERIAL

#### 3.1. Descripción de la zona de estudio

El área de estudio está situada en la zona centro sur de Navarra, entre la Zona Media y la Ribera Alta y parte de Tierra Estella y Valdizarbe. En este entorno se define una zona rectangular comprendida entre los municipios de Miranda de Arga, Tafalla y San Martín de Unx por la parte norte, Ujué, Pitillas, Santaclara y Mélida por el este, Murillo el Cuende, Caparrosa, Marcilla, Funes y Azagra por el sur, y San Adrián, Andosilla y Lerín por el oeste (**Figura 3**), abarcando una superficie aproximada de 73.060 ha.



**Figura 3.** Situación de la zona de estudio.

En cuanto a la climatología, la zona de estudio queda emplazada en la denominada *Zona sur* de Navarra, caracterizada por precipitaciones más bien escasas y veranos cálidos y secos. Las precipitaciones anuales varían entre los 500 y 600 l/m<sup>2</sup>, con veranos secos (entre 65 y 115 l/m<sup>2</sup>). En cuanto a las temperaturas, la media anual varía entre 13,5 y 14,4°C de norte a sur, alcanzando temperaturas máximas medias de hasta 31°C en agosto y 10°C en enero.



Al ser una zona situada en la depresión del Ebro, presenta un relieve definido por áreas llanas, principalmente terrazas cuaternarias, cruzadas por pequeñas lomas y sierras. Los valles siguen los grandes ríos que atraviesan la zona (Arga, Aragón y Cidacos), definidos por suaves cambios en el relieve, con pendientes de entre el 5 y el 10 %.

La principal ocupación del suelo en la zona de estudio es agrícola (73%), seguido del uso forestal (18%), improductivo (5%) y otros usos (4%). Dentro de las cubiertas agrícolas, los cultivos herbáceos en secano (principalmente cebada y trigo blando) son los que presentan mayor ocupación de la zona (51%), seguido de cultivos herbáceos en regadío (maíz, trigo blando, cebada y girasol esencialmente) (39%) y cultivos leñosos en regadío (viñedo vinificación en su gran mayoría) (8%). La **Tabla 3** muestra los calendarios agrícolas de los principales cultivos, correspondiendo el sombreado verde claro a la siembra, el verde oscuro a la germinación y al periodo vegetativo, el rosa claro al período reproductivo, el rosa oscuro al llenado de fruto y madurez, y el amarillo a la cosecha.

**Tabla 3.** Calendario de cultivos principales de la zona de estudio divididos en: cultivos de invierno (cebada y trigo blando), de verano (girasol y maíz) y permanente (viñedo vinificación – Vid- ).

	Sep.	Oct.	Nov.	Dic.	Ene.	Feb.	Mar.	Abr.	May.	Jun.	Jul.	Ago.
Cebada												
Trigo blando												
Girasol												
Maíz												
Vid												
Siembra, germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.												

### 3.2. Datos de partida

Como datos de partida se contó con imágenes capturadas por los satélites Sentinel-1 y Sentinel-2 sobre Navarra, para un período de estudio de 01/09/2018 a 31/12/2019, coincidiendo con la campaña agrícola del año 2019. Dichas imágenes ya habían sido utilizadas en estudios previos, teniendo estas, por lo tanto, cierto nivel de procesamiento, comentado en los siguientes apartados.

#### 3.2.1. Imágenes Sentinel-1

Las imágenes Sentinel-1 proporcionadas fueron capturadas en tres pasadas diferentes: una ascendente (ASC\_103) y dos descendentes (DESC\_8 y DESC\_81) (**Figura 4**), en modo *Interferometric Wide Swath* (IW), con todas las adquisiciones de los satélites gemelos S-1A y S-1B teniendo, por lo tanto, imágenes con una resolución temporal de 6 días.



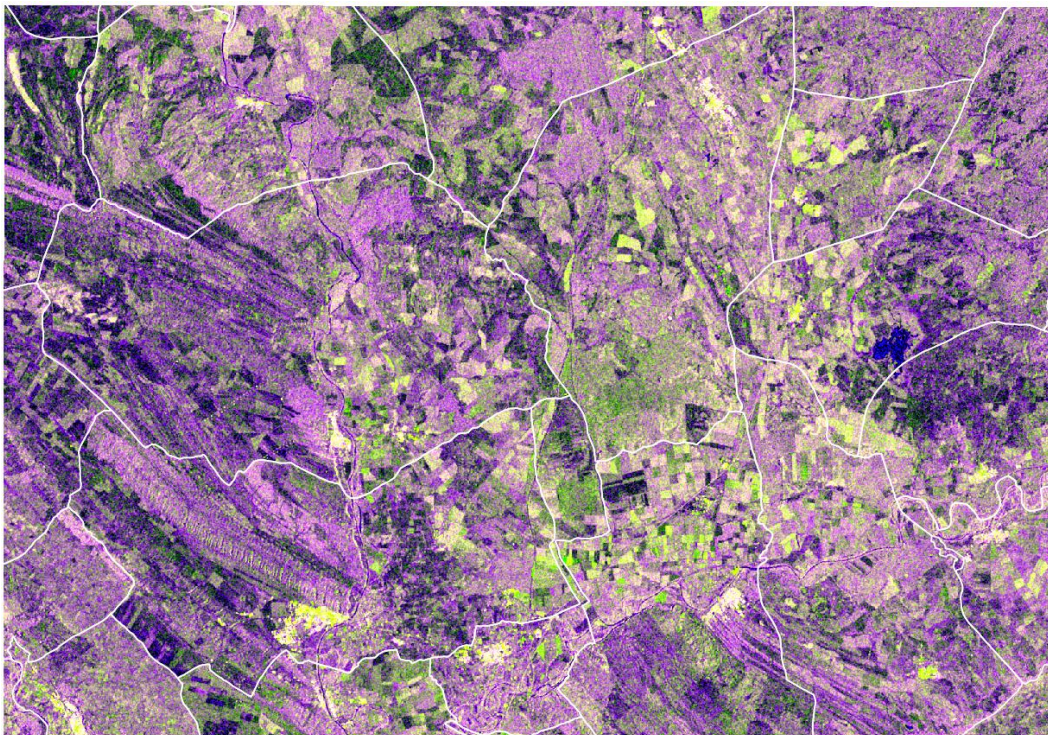
**Figura 4.** Huella satelital (zona sombreada en rojo) de la pasada ASC\_103 (izquierda), la DESC\_81 (centro) y la DESC\_8 (derecha) sobre la zona de estudio (zona sombreada naranja).

Las imágenes fueron descargadas en nivel *Ground Range Detected* (GRD), en polarización cruzada (VH) y directa (VV) desde la página web *Copernicus Open Access Hub* de forma gratuita. Como se ha comentado con anterioridad, dichas imágenes radar fueron utilizadas en estudios previos, recibiendo el siguiente procesado:

- 1) Eliminación del ruido térmico: consiste en eliminar el ruido emitido por el propio satélite debido a su temperatura. Es un efecto que se manifiesta más en zonas de retrodispersión baja, como el asfalto. Las imágenes descargadas no tenían dicha corrección aplicada (la información necesaria para conocer si el ruido térmico ha sido corregido se encuentra en los metadatos asociados a la imagen), por lo que se realizaron las correcciones para ambas polarizaciones (VV y VH).
- 2) Corrección de la órbita: para poder garantizar una corrección geométrica precisa, es necesario sustituir los datos orbitales aproximados del producto radar por los datos reprocesados disponibles días o semanas después de la generación del producto.
- 3) Calibración a  $\beta^0$ : consiste en transformar las unidades de las imágenes de niveles digitales a valores de coeficientes de retrodispersión  $\beta^0$ , con el fin de poder diferenciar los diferentes mecanismos de retrodispersión producidos en las imágenes según las características del terreno, además de ser un requisito indispensable para poder comparar imágenes adquiridas por el mismo sensor en diferentes fechas.
- 4) Filtrado del moteado: para eliminar el efecto moteado (*speckle*) del que se ven afectadas las imágenes radar que introduce variabilidad en los valores de los píxeles, generando un aspecto de *sal y pimienta*. Este efecto reduce la calidad radiométrica, dificultando su interpretación, por ello es necesario aplicar filtros que eliminen dicho moteado sin modificar la estructura de la imagen. A las imágenes utilizadas en el presente trabajo se les aplicó el filtro adaptativo Gamma Map 3x3 (Lopes et al., 1990), enfocado en mantener las texturas de las imágenes, en las bandas de las dos polarizaciones.

- 5) Corrección topográfica de  $\beta^0$  a  $\gamma_{TF}^0$ : consiste en corregir las variaciones radiométricas producidas a la posición topográfica de los píxeles, donde la retrodispersión aumenta cuando la ladera está orientada al sensor, y viceversa. Este efecto se corrige aplicando el método *Terrain Flatening* (Small, 2011), a partir del cual se obtienen datos de coeficiente de retrodispersión normalizados topográficamente  $\gamma_{TF}^0$  en unidades de gamma0
- 6) Ortorrectificación: proceso mediante el cual se corrigen las desviaciones geométricas de las imágenes generadas por la superficie de la tierra, así como por las distorsiones del sensor. En este caso, el modelo utilizado es el *Range-Doppler*, un modelo riguroso mediante el cual se resuelve matemáticamente las ecuaciones del rango y del efecto Doppler (Small & Schubert, 2008), con el fin de obtener posiciones precisas de cada punto de las imágenes.

Para cada fecha de captura se obtuvieron, por lo tanto, cuatro bandas, correspondientes al ángulo de incidencia local (B1, banda 1), coeficiente de retrodispersión cruzada VH en unidades de gamma0 ( $\gamma_{VH}^0$ ) en decibelios (dB) (B2), coeficiente de retrodispersión directa VV en unidades gamma0 ( $\gamma_{VV}^0$ ) en dB (B3) y una cuarta banda B4 con el ratio entre el  $\gamma_{VH}^0$  y  $\gamma_{VV}^0$  en dB. La **Figura 5** muestra una composición en RGB (B2, B3, B4) de la zona de estudio.

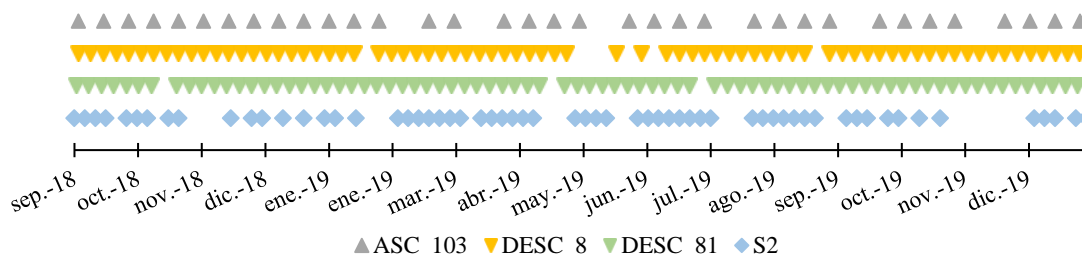


**Figura 5.** Composición en RGB (VV, VH, VH/VV) de imágenes adquiridas por el satélite Sentinel-1 de la zona de estudio para la fecha de captura del 08/09/2018 de la pasada DESC\_81.

Los momentos de captura de las imágenes Sentinel-1 y Sentinel-2 disponibles para cada fecha quedan representados en la **Figura 6**. Como puede observarse, se obtienen un total de 35 imágenes de la pasada ASC\_103, 74 imágenes de la pasada DESC\_8 y 78 de la pasada DESC\_81. En la



pasada ascendente 103 el número de imágenes disponibles es mucho menor que en las otras pasadas debido a que el satélite S-1A no capturó imágenes sobre la zona de estudio durante el periodo de interés.



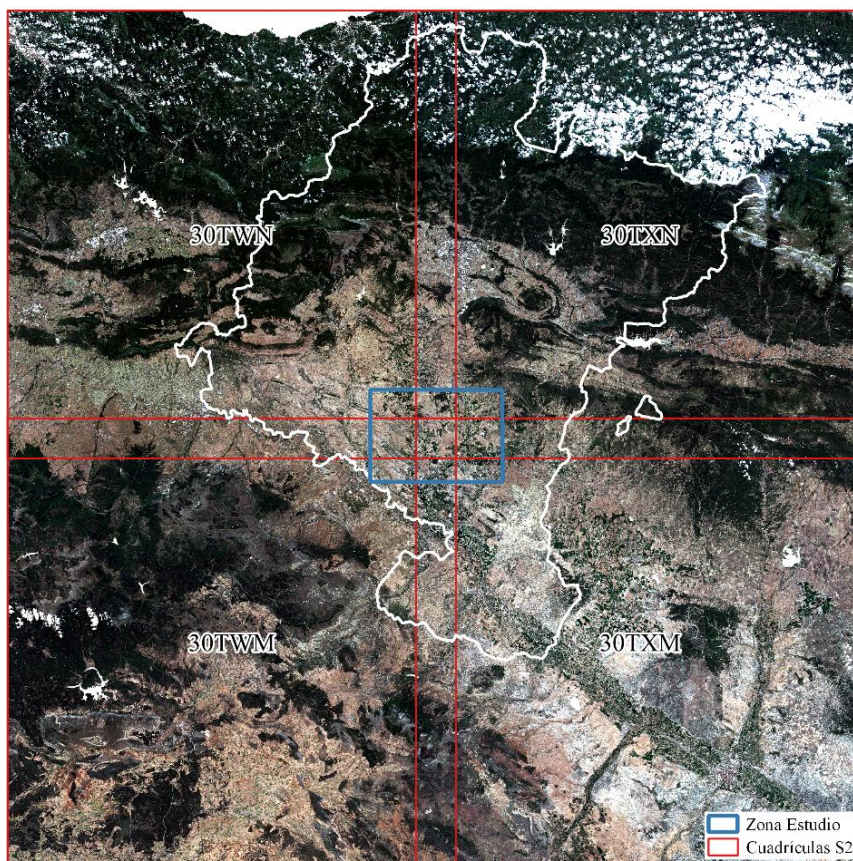
**Figura 6.** Fechas de adquisición de imágenes Sentinel-1 (ASC\_103, DESC\_8 y DESC\_81) y Sentinel-2 (S2).

### 3.2.2. Imágenes Sentinel-2

Las imágenes Sentinel-2 proporcionadas fueron descargadas en nivel L2A (reflectancias en superficie), y se les aplicó un proceso de mosaico para unir las cuatro cuadrículas S-2 en las que está comprendida la zona de estudio: 30TXN, 30TXM, 30TWN y 30TWM (**Figura 7**), un remuestreo a 10 m (debido a la diferencia en la resolución espacial de las bandas espectrales de S-2) y una máscara de nubes (obtenida mediante el algoritmo SCL de Sen2Cor) para descartar píxeles contaminados con nubes o sombras de nubes. De la misma manera que Sentinel-1, las imágenes Sentinel-2 obtenidas fueron adquiridas por los satélites gemelos S-2A y S-2B, permitiendo tener imágenes con una resolución temporal de 5 días, aunque se descartaron imágenes con una cobertura de nubes igual o superior al 60% de la superficie de la zona de estudio. Para cada fecha de captura se facilitaron las bandas B2 (azul), B3 (verde), B4 (rojo), B5 (red edge 1), B6 (red edge 2), B7 (red edge 3), B8 (NIR 1), B8A (NIR 2), B11 (SWIR 1) y B12 (SWIR 2) de Sentinel-2 (**Figura 1** y **Tabla 4**). En la **Figura 6** puede observarse la disponibilidad de las 61 imágenes Sentinel-2 proporcionadas a lo largo del periodo de estudio.

**Tabla 4.** Especificaciones de las bandas espectrales de Sentinel-2.

Sentinel 2, MSI			
Resolución espacial	10 m	20 m	60 m
Bandas	B2 (azul): 458 - 522 nm B3 (verde): 543 - 577 nm B4 (rojo): 650 - 680 nm B8 (NIR): 785 - 0.900 nm	B5 (Vegetation Red Edge): 698 - 713 nm B6 (Vegetation Red Edge): 733 - 748 nm B7 (Vegetation Red Edge): 773 - 793 nm B8A (Vegetation NIR): 855 - 875 nm B11 (SWIR): 1.565 - 1.655 nm B12 (SWIR): 2.100 - 2.280 nm	B1 (Coastal Aerosol): 433 - 453 nm B9 (Water Vapour): 935 - 955 nm B10 (SWIR-Cirrus): 1.365 - 1.395 nm
Resolución radiométrica	12 bits		



**Figura 7.** Composición RGB color natural de las adquisiciones Satélite-2 el día 01/09/2019 mostrando las cuatro cuadrículas que comprenden la zona de estudio.

### 3.2.3. Parcelas agrícolas de la zona de estudio

La *Sección de Inspecciones del Servicio del Organismo Pagador del Departamento de Desarrollo Rural y Medio Ambiente (DRyMA)* del Gobierno de Navarra genera un archivo vectorial con la información relativa a las declaraciones PAC presentados por los agricultores en cada campaña agrícola. Para el presente trabajo se utilizó el archivo vectorial anonimizado en formato *shape* de la campaña agrícola de 2019 recortado a la zona de estudio. El archivo cuenta con diferentes campos que permiten identificar la parcela como son el número identificador de cada parcela, el municipio, polígono, parcela y recinto, el cultivo declarado o la superficie cultivada en m<sup>2</sup>. La zona de estudio queda cubierta por 27.954 parcelas agrícolas, cultivadas con 95 tipos de cultivos tanto en secano como en regadío, abarcando una superficie cultivada de 48.229 ha.

Debido a la gran cantidad de cultivos de partida, se decidió reducir el número de cultivos con los que trabajar, escogiendo aquellos que fueran preponderantes y, por tanto, representativos de la zona de estudio. Dicha selección se realiza con el objetivo doble de evitar la inclusión de cultivos minoritarios con una muestra (número de parcelas) insuficiente para el análisis y para reducir el coste computacional y magnitud de trabajo. La selección de cultivos (**Tabla 5**) queda definida en el apartado 4.1.

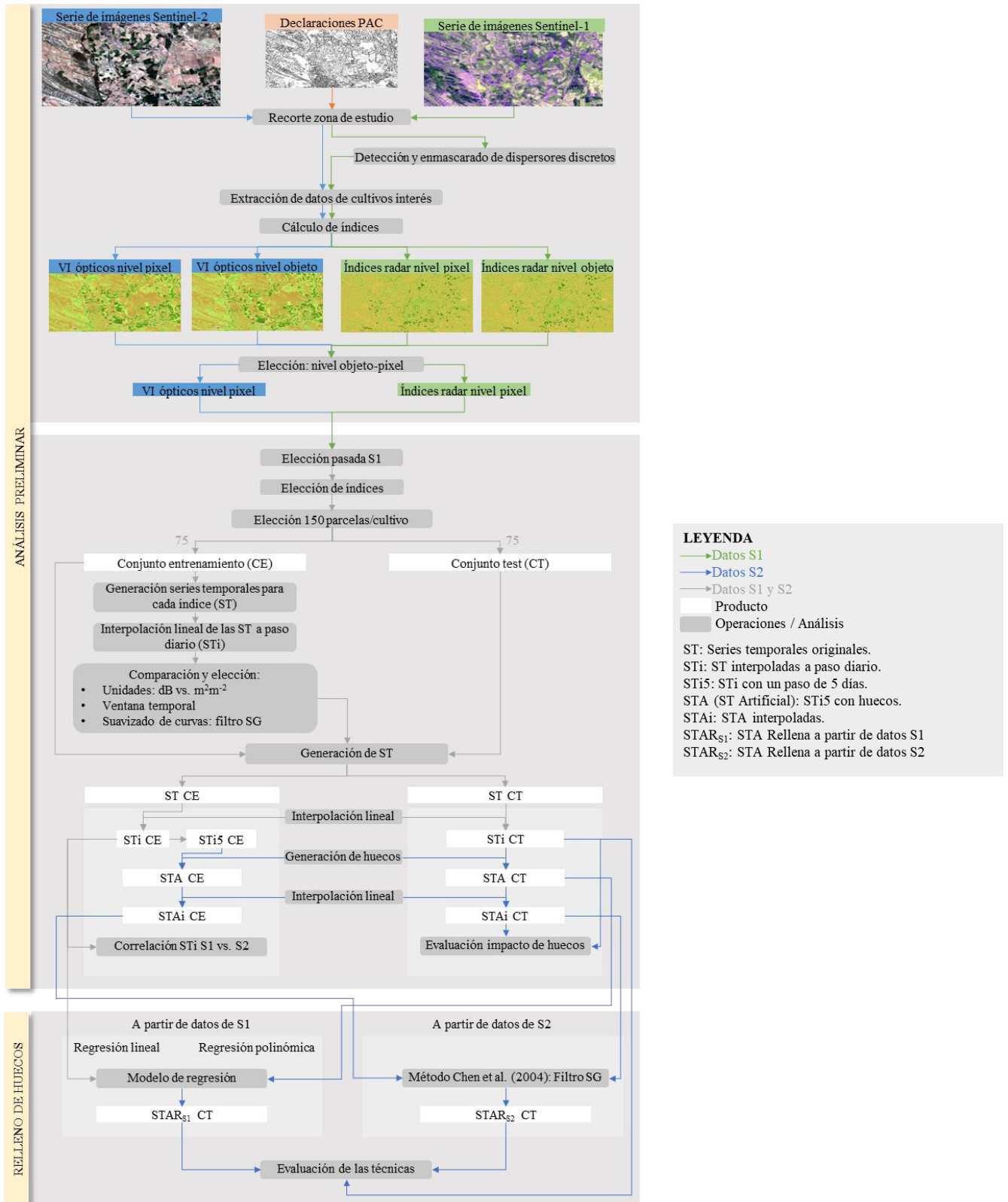
**Tabla 5.** Número de parcelas y superficie ocupada por los cultivos de interés.

	<b>Cultivo</b>	<b>nº parcelas</b>	<b>Superficie (ha)</b>
CULTIVOS INVIERNO	CEBADA SECANO	1.315	6.550
	CEBADA REGADÍO	587	1.912
	TRIGO BLANDO SECANO	782	3.363
	TRIGO BLANDO REGADÍO	678	2.780
CULTIVOS VERANO	GIRASOL	151	892
	MAÍZ	865	4.653
C. PERMANENTE	VIÑEDO VINIFICACIÓN	1.098	2.849

Además, en el ANEXO I. INFORMACIÓN DE LOS 95 CULTIVOS DE PARTIDA se aporta información detallada de los 95 cultivos de partida.

## 4. METODOLOGÍA

La **Figura 8** muestra un esquema general, a modo de diagrama de flujo, de la metodología aplicada, detallada en los siguientes apartados.



**Figura 8.** Diagrama de flujo de la metodología aplicada en el trabajo.



#### 4.1. Análisis preliminar

El primer paso fue recortar las imágenes Sentinel-1 y Sentinel-2, a partir del archivo vectorial de la zona de estudio.

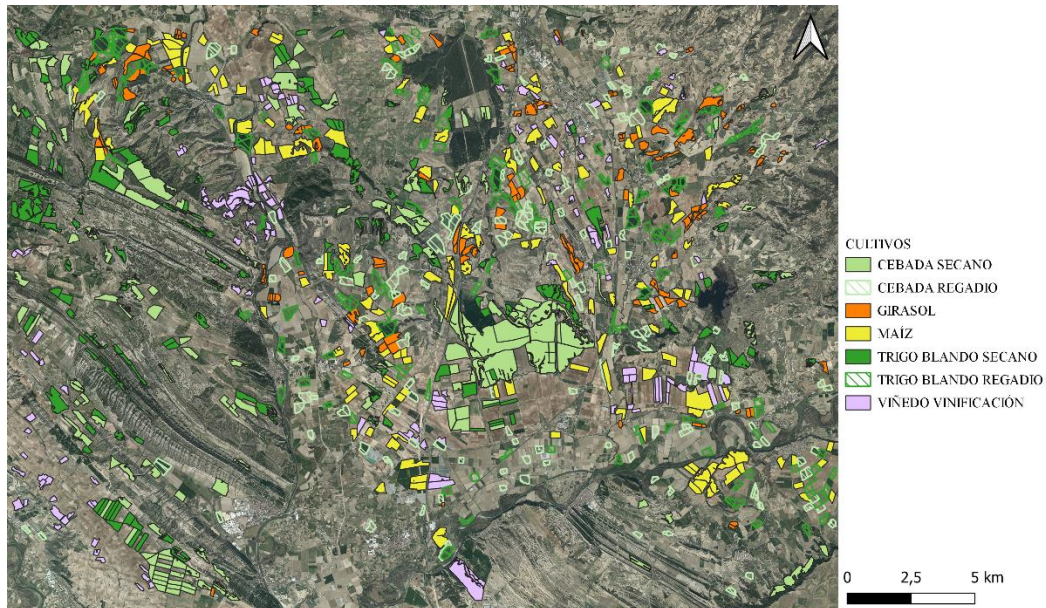
A partir de estas imágenes recortadas, se detectaron y enmascararon aquellos dispersores discretos que pudieran provocar alteraciones en los valores de retrodispersión de los cultivos por fenómenos de doble rebote. Estos fenómenos son muy improbables en los cultivos presentes en el área de estudio para la longitud de onda de Sentinel-1 (5,405 GHz). Sin embargo, elementos puntuales (edificios u otras estructuras, **Figura 9**) presentes en las inmediaciones de las parcelas agrícolas pueden generar este tipo de fenómenos e incluso alterar los valores de los píxeles vecinos, impidiendo una correcta interpretación de los resultados. Por tanto, los píxeles con  $\gamma_{VH}^0$  superior a  $0,32 \text{ m}^2\text{m}^{-2}$  y/o  $\gamma_{VV}^0$  superior a  $1 \text{ m}^2\text{m}^{-2}$  fueron excluidos del análisis (marcados como NaN, *Not a Number*, en Python).



**Figura 9.** Ejemplo de dos reflectores puntuales en la zona de estudio: (a) aerogenerador en el borde de una parcela y (b) edificio.

Posteriormente, previo al cálculo de los índices, se realizó la selección de cultivos de interés. Para ello, primero se eligieron aquellos cultivos que tuvieran una representación mínima de 100 parcelas (teniendo en cuenta únicamente aquellas parcelas cuya superficie fuera igual o superior a 1 ha). Una vez realizada esta primera selección, se excluyeron aquellas parcelas que, durante el periodo de estudio, tuvieran un porcentaje medio de píxeles cubiertos por nube superior al 50% (se consideraron píxeles cubiertos por nube aquellos que no tuvieran dato, resultado de la aplicación de la máscara de nubes y sombras), eliminando así valores de reflectividad que pudieran ‘contaminar’ los resultados e interpretaciones. Tras esta discriminación, se eligieron aquellos cultivos con mayor representación en la zona (aquellos con mayor número de parcelas), diferenciando entre cultivos de invierno (cebada y trigo blando, haciendo también distinción entre secano y regadío), cultivos de verano (girasol y maíz) y un cultivo permanente (viñedo).

vinificación), cuya información queda detallada en la **Tabla 5** y su distribución representada en la **Figura 10**. Distribución de las parcelas seleccionadas en la zona de estudio..



**Figura 10.** Distribución de las parcelas seleccionadas en la zona de estudio.

Una vez seleccionados los cultivos y las parcelas de interés, se procedió a calcular diferentes índices ópticos y radar, **Tabla 6** y **Tabla 7**, seleccionados de la literatura científica por su amplio uso y aplicación en el ámbito agronómico (ver apartado 1.2.1).

**Tabla 6.** Definiciones y fórmulas de los índices de vegetación ópticos calculados.

Índice	Abreviación	Fórmula	Referencia
Normalised Difference Vegetation Index	NDVI	$\frac{\rho_{B8} - \rho_{B4}}{\rho_{B8} + \rho_{B4}}$	Rouse et al. (1973)
Soil Adjusted Vegetation Index	SAVI	$\frac{\rho_{B8} - \rho_{B4}}{\rho_{B8} + \rho_{B4} + L} (1 + L), L = 0,5$	Huete (1988)
Renormalized Difference Vegetation Index	RDVI	$\frac{\rho_{B8} - \rho_{B4}}{\sqrt{\rho_{B8} + \rho_{B4}}}$	Roujean & Breon (1995)
Normalized Difference Moisture Index	NDMI	$\frac{\rho_{B8} - \rho_{B12}}{\rho_{B8} + \rho_{B12}}$	Cibula et al. (1992)
Specific Leaf Area Vegetation Index	SLAVI	$\frac{\rho_{B8}}{\rho_{B11} + \rho_{B4}}$	Lymburner et al. (2000)
Chlorophyll index Red-Edge $CI_{RE}$	$CI - RE$	$\frac{\rho_{B7}}{\rho_{B5}} - 1$	Gitelson, Gritz, et al. (2003)
Modified Chlorophyll Absorption in Ratio Index	MCARI	$[(\rho_{B5} - \rho_{B4}) - 0,2(\rho_{B5} - \rho_{B3})](\rho_{B5}/\rho_{B4})$	Daughtry et al. (2000)
Meris Terrestrial Chlorophyll Index	MTCI	$\frac{\rho_{B6} - \rho_{B5}}{\rho_{B5} - \rho_{B4}}$	Dash & Curran (2010)
Red-edge NDVI	$NDVI_{re}$	$\frac{\rho_{B7} - \rho_{B6}}{\rho_{B7} + \rho_{B6}}$	Gitelson & Merzlyak (1994)
Sentinel-2 LAI index	SeLI	$\frac{\rho_{B8} - \rho_{B5}}{\rho_{B8} + \rho_{B5}}$	Pasqualotto et al. (2019)

**Tabla 7.** Definiciones y fórmulas de los índices radar calculados.

Índice	Abreviación	Fórmula	Referencia
Coef. retrodispersión cruzada	VH	$\gamma_{VH}^0$	
Ceof. retrodispersión directa	VV	$\gamma_{VV}^0$	
Suma VV, VH	RSUM	$\gamma_{VV}^0 + \gamma_{VH}^0$	
Diferencia VV, VH	RDIF	$\gamma_{VV}^0 - \gamma_{VH}^0$	
Producto VV, VH	RMULTI	$\gamma_{VH}^0 \times \gamma_{VV}^0$	
Ratio VH/VV	RDIV	$\frac{\gamma_{VH}^0}{\gamma_{VV}^0}$	
<i>Radar Vegetation Index</i>	RVI <sub>VV</sub>	$\frac{4\gamma_{VH}^0}{\gamma_{VV}^0 + \gamma_{VH}^0}$	Nasirzadehdizaji et al. (2019)
<i>Vertical Dual de-polarization Index</i>	VDDPI	$\frac{\gamma_{VV}^0 + \gamma_{HH}^0}{\gamma_{VV}^0}$	Periasamy (2018)
<i>Inverse Dual-Pol Diagonal Distance</i>	IDPDD	$\frac{(\gamma_{VV(max)}^0 - \gamma_{VV(i)}^0) + \gamma_{VH(i)}^0}{\sqrt{2}}$	Periasamy (2018)

Este cálculo se realizó de dos formas: a nivel de pixel y agrupando los píxeles a nivel de objeto (parcelas agrícolas) con el fin de analizar si, mediante la segunda opción, se obtenían resultados aproximados a la primera. El interés de analizar ambas opciones fue debido al ahorro en almacenamiento y al coste computacional que supondría poder guardar y ejecutar la información a nivel de parcela. Para ello, se calculó y analizó el NRMSE (raíz del error cuadrático medio normalizado), la correlación y el bias entre el cálculo por pixel y por parcela.

A partir de los índices ópticos y radar calculados, el primer paso en el tratamiento de datos fue elegir para el relleno de huecos una de las tres pasadas de Sentinel-1 disponibles. Para ello, se valoraron para cada pasada dos aspectos, por un lado, el número de imágenes disponibles en la serie temporal de estudio (**Figura 6**), y, por otro, la correlación entre los índices de S1 y S2 para todo el conjunto de datos. Estos valores de correlación medios se obtuvieron calculando, primero, la matriz de correlación entre los índices de Sentinel-1 y Sentinel-2 por cultivo y por parcela, y, promediando, posteriormente, dichos valores. Se escogió la pasada que presentó una combinación de los dos parámetros superior. Los satélites S1 y S2 capturan imágenes en días diferentes, por lo que, para poder realizar correlaciones entre los diversos índices, fue necesario disponer de ambos datos en las mismas fechas, para lo que se realizó previamente una interpolación temporal a datos diarios.

El siguiente paso fue reducir el número de índices a considerar en el análisis, agrupando aquellos que aportaran información similar. Para realizar dichos grupos, se volvieron a analizar las matrices de correlación previamente mencionadas, identificando los índices muy correlacionados

entre sí y eliminando aquellos que presentaran información redundante. Se redujo el número de índices de Sentinel-2 a cinco y de Sentinel-1 a cuatro (ver apartado 5.3).

Una vez decididos los cultivos, las pasadas y los índices de trabajo, se seleccionó una muestra equilibrada de parcelas por cultivo. Se estableció un tamaño de muestra de 150 parcelas por cultivo, cifra limitada por el cultivo con menor número de parcelas, en este caso el girasol (**Tabla 5**). La selección de parcelas se realizó de forma intencional (no aleatoria) escogiendo aquellas con mayor superficie y menor porcentaje de píxeles cubiertos por nube. La muestra se dividió aleatoriamente en dos conjuntos de datos aplicando el método *Hold-out*: un conjunto de entrenamiento (CE) con el 50% de los datos (75 parcelas) y un conjunto de test (CT) con el 50% restante (75 parcelas).

Sobre el CE se realizaron una serie de análisis con el fin de obtener las series temporales óptimas para cada índice:

- Unidades Sentinel-1: se analizó en qué unidades resultaba óptimo el cálculo de los índices Sentinel-1 comparando dos alternativas:
  - Unidades de retrodispersión en decibelios (dB)
  - Unidades lineales ( $m^2m^{-2}$ )
- Ventana temporal (VT) de análisis: se evaluó qué ventana temporal resultaba idónea para el análisis de la serie temporal de cada cultivo comparando dos alternativas:
  - Todos los datos disponibles desde 01/09/2018 hasta 31/12/2019 ( $ST_C$ )
  - Ventana temporal con el ciclo fenológico del cultivo diferenciando cultivos de verano, invierno y permanentes ( $ST_{VT}$ ).
- Suavizado de la serie temporal: algunos índices pueden tener una excesiva variabilidad temporal, debido a otros factores distintos al desarrollo del cultivo, por ello, se evaluó si un suavizado de la serie temporal mejoraba los resultados, comprobando:
  - Curva con datos sin suavizar ( $ST_{VT}$ )
  - Curva con datos suavizados mediante el filtro de Savitzky-Golay ( $ST_{VT,SG}$ ).

Para realizar esta comparativa, se reprodujeron las matrices de correlación entre índices de Sentinel-1 y Sentinel-2, para las tres opciones planteadas, eligiendo aquellas que obtuvieran una mayor correlación entre índices de S1 y S2, generando así las series temporales finales para los dos conjuntos de datos CE y CT, sobre las que se evalúan en el apartado siguiente las diferentes técnicas de relleno de huecos.

Las ventanas temporales de cada cultivo (**Tabla 8**) se establecieron en función de su ciclo fenológico (**Tabla 3**).



**Tabla 8.** Ventanas temporales aplicadas en la serie temporal de estudio para cada cultivo.

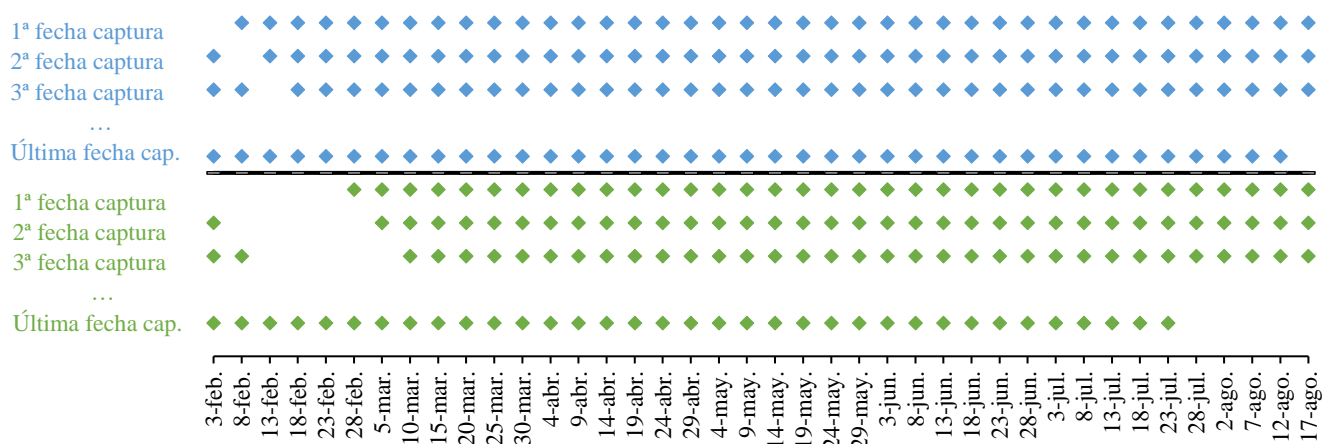
VENTANA TEMPORAL	CULTIVOS
2019/02/01 – 2019/07/31	CEBADA, TRIGO
2019/04/01 – 2019/10/15	GIRASOL
2019/06/01 – 2019/12/31	MAÍZ

En cuanto al suavizado de las series temporales, se decidió aplicar el filtro de Savitzky-Golay (SG) por haber sido ampliamente utilizando en este tipo de análisis (Joshi et al., 2016; Kandasamy et al., 2013; Liu et al., 2017). Este filtro tiene como base el cálculo de una regresión polinómica local que determina el nuevo valor de cada punto de la serie. Los parámetros que definen el filtro son el tamaño de la ventana de suavizado (cuanto menor sea este valor, más se ajustará la nueva serie suavizada a la original, obteniendo un suavizado menor, y viceversa) y el grado del polinomio (al contrario del tamaño de la ventana, cuanto mayor sea el grado del polinomio, mayor será el ajuste de la función a los datos originales, obteniendo un suavizado menor, y viceversa). El objetivo de aplicar dicho filtro fue obtener curvas que representaran las variaciones de los diferentes índices (tanto radar como ópticos) a lo largo del ciclo fenológico de cada cultivo, eliminando el ruido que pudiera deberse a la eventual presencia de nubes y a variabilidad atmosférica, en el caso de Sentinel-2 y a la influencia de otros factores en la retrodispersión (humedad del suelo, rugosidad, eventos meteorológicos, etc.) en el caso de Sentinel-1. En este caso se comparó el suavizado considerando dos tamaños de ventana diferentes (15 y 31 días) y tras varias pruebas se seleccionó un grado de polinomio de 1.

Finalmente, antes de comenzar con la fase de relleno de huecos, se realizaron dos últimos análisis sobre las series temporales de Sentinel-2. Por un lado, utilizando el conjunto de datos de entrenamiento (CE), se realizó un último análisis sobre la correlación entre los índices de S1 y S2 utilizando las series temporales derivadas del procesamiento definitivo (series temporales interpoladas, suavizadas mediante el filtro SG, con una ventana temporal para cada cultivo en función de su fenología). Por otro lado, se utilizó el conjunto de test (CT) para evaluar el impacto de los huecos en las series temporales de índices de Sentinel-2. Para ello, se realizó lo siguiente

- Interpolación lineal: se realizó una primera interpolación lineal de las series temporales de Sentinel-2, rellenando de esta manera los posibles huecos que tuviera la serie original causados por la presencia de nubes. Productos finales: series temporales interpoladas a paso diario (STi) y series temporales interpoladas con un paso de 5 días (STi5).
- Simulación de huecos: se generaron huecos en las STi5, simulando los huecos que podrían haberse generado de forma natural (en fechas de captura de imagen). Esta introducción de huecos no se produjo de manera aleatoria, sino que, para cada cultivo, se generaron huecos con duraciones determinadas: huecos de una captura (1c), de dos (2c), tres (3c), cuatro (4c), cinco (5c), diez (10c), quince (15c) y veinte (20c) capturas seguidas,

en diferentes momentos del ciclo del cultivo (se fueron generando los huecos a lo largo de todas las capturas de la serie temporal). Trabajar con simulaciones permite comparar el dato rellenado mediante las técnicas propuestas con el dato realmente captado. La **Figura 11** muestran varios ejemplos de cómo se introdujeron los huecos con duraciones de 1c (azul) y 5c (verde) sobre una parcela de cebada, en diferentes momentos del ciclo (a partir de la primera, segunda, tercera y última captura posible de la serie temporal). Producto final: series temporales artificiales (STA).



**Figura 11.** Ejemplo de huecos de una y cinco capturas generados en una serie temporal, a partir de la primera, segunda, tercera y última fecha de adquisición.

- c) Interpolación lineal final: para simular la serie temporal interpolada que se obtendría en el caso de una serie con huecos, se realizó una interpolación lineal final a partir de las STA, rellenando los huecos generados en el punto anterior (b). Producto final: series temporales artificiales interpoladas (STAi).

Como se ha comentado previamente, se calculó la correlación entre los índices de Sentinel-1 y Sentinel-2 a partir de las STi del CE generadas en el punto (a). Por otro lado, se evaluó el impacto que tienen los huecos en las series temporales de Sentinel-2, en función del momento y la duración de estos, utilizando como indicador el RMSE entre los datos rellenados y los originales (calculado únicamente para las fechas dónde se simularon huecos y no para toda la serie temporal).

## 4.2. Aplicación de las técnicas de relleno de huecos

Una vez finalizado el tratamiento de datos, el siguiente paso consistió en aplicar diferentes técnicas de relleno sobre las STA del CT. Con el fin de acotar el trabajo para esta fase se consideró únicamente el índice NDVI, aplicando las técnicas de relleno de huecos que se detallan a continuación.

### 4.2.1. Relleno a partir de datos de Sentinel-1

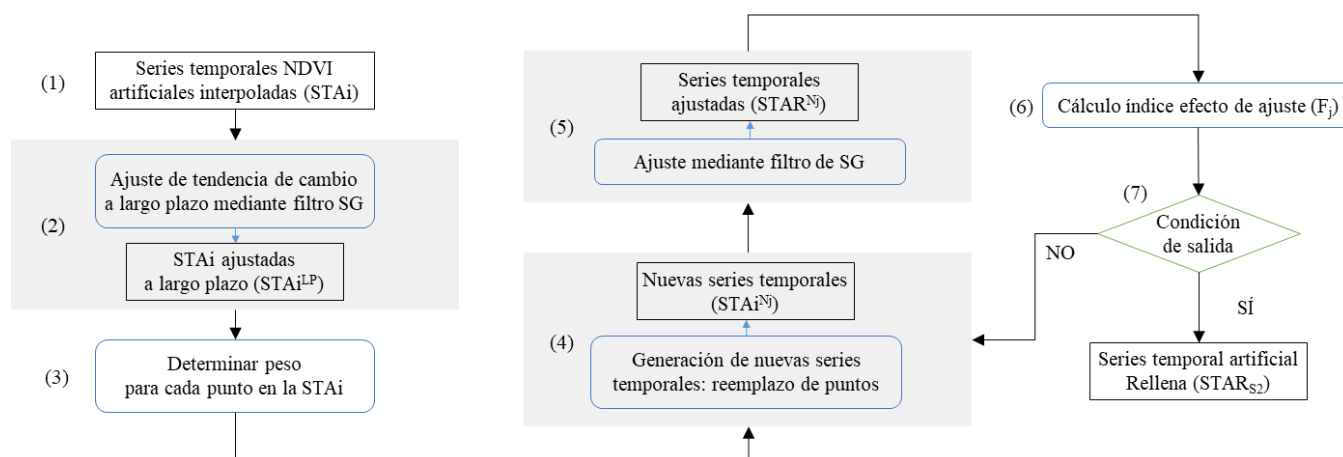
La reconstrucción de datos de NDVI a partir de índices de Sentinel-1 se realizó mediante la aplicación de regresiones lineales y polinómicas ajustadas con el conjunto de datos de

entrenamiento (CE). En el caso de las regresiones polinómicas, se evaluaron los resultados con polinomios de distinto grado obteniendo resultados óptimos con segundo grado (en cuanto a la correlación y coste computacional). Para ello, se entrenaron los modelos de regresión con las STi del CE para, posteriormente, aplicarlos sobre las STA del CT, planteando un problema univariante (reconstruir los huecos de NDVI a partir de los cuatro índices de Sentinel-1 de manera individual). A partir del modelo entrenado, se obtuvieron cuatro series de NDVI que se consideraron de *referencia*, de las que se extrajeron los valores de NDVI puntuales con los que se rellenaron los huecos de las series de NDVI ( $STAR_{S1}$ , serie temporal artificial reconstruida a partir de índices S1). Estas operaciones se realizaron mediante la librería de aprendizaje automático de software libre para el lenguaje de programación Python *Scikit-learn*.

Para cuantificar la calidad de la técnica de relleno, se calculó el RMSE entre los valores de NDVI que coincidieran con las fechas en las que se hubiera reconstruido los huecos de la nueva serie temporal reconstruida ( $STAR_{S1}$ ) y de la serie temporal interpolada (STi) del CT.

#### 4.2.2. Relleno a partir de datos de la propia serie de Sentinel-2

Para reconstruir los huecos en la serie de NDVI a partir de datos de la propia serie se aplicó el método planteado por Chen et al. (2004), basado en la aplicación reiterada del filtro de Savitzky-Golay sobre la serie temporal inicial, con el fin de obtener una serie temporal final de alta calidad. Para ello, el método plantea el procedimiento (**Figura 12**) detallado a continuación (aplicado tanto para el CE, a partir del cual se extrajeron los parámetros del filtro de SG que posteriormente se aplicaron sobre el CT, como para el CT, sobre el que se validaron los resultados):



**Figura 12.** Diagrama de flujo del método planteado por Chen et al. (2004).

- 1) Interpolación lineal de valores de NDVI con nube: en este caso se partió de las series temporales artificiales interpoladas (STAi) obtenidas en el apartado 4.1.
- 2) Ajuste de tendencia de cambio a largo plazo mediante filtro de Savitzky-Golay ( $STAi^{LP}$ ): con este ajuste se pretende generar una nueva serie temporal de NDVI en la que se eviten

aquellos valores que puedan haber sido afectados por nubes o malas condiciones atmosféricas. Es decir, se pretende obtener una serie temporal que represente los cambios en el ciclo vegetativo de un cultivo. Dicho ajuste se realiza suavizando la STi mediante la aplicación del filtro de SG. Estos parámetros se obtuvieron a partir del CE, definiendo un tamaño de ventana de 19 y un polinomio de grado 1.

- 3) Determinación del peso para cada punto de la serie temporal de NDVI: la serie temporal generada en el paso anterior ( $STAi^{LP}$ ) se utilizó para determinar el peso de cada punto de la  $STAi$ , con el fin de poder definir un índice que mida el efecto del ajuste y obtener una serie temporal final reconstruida de alta calidad que represente las variaciones de NDVI durante el ciclo de la vegetación. Estos pesos ( $W_i$ ) se calculan a partir de la siguiente función:

$$W_i = \begin{cases} 1 & NDVI_{i,o} \geq NDVI_{i,LP} \\ 1 - d_i/d_{max} & NDVI_{i,o} < NDVI_{i,LP} \end{cases}$$

donde  $NDVI_{i,o}$  y  $NDVI_{i,LP}$  corresponden a valores de NDVI para una fecha de captura  $i$  en la serie original y en la serie ajustada a largo plazo, respectivamente,  $d_i = |NDVI_o - NDVI_{LP}|$  y  $d_{max}$  es el máximo de los  $d_i$ .

- 4) Generación de nuevas series temporales de NDVI ( $STAi^{Nj}$ ): Chen et al. (2004) observaron que la  $STAi^{LP}$  presentaba valores de NDVI superiores que la  $STAi$  en puntos donde hubiera ruido, sugiriendo un reemplazo de los puntos ruidosos de la  $STAi$  por los valores correspondientes de la  $STAi^{LP}$ , para reajustar de nuevo la serie temporal, acercándose así a la envolvente superior de datos originales del NDVI. Este ajuste a la envolvente se realizó aplicando la siguiente fórmula:

$$NDVI_i^{Nj} = \begin{cases} NDVI_{i,o} & NDVI_{i,o} \geq NDVI_{i,LP} \\ NDVI_{i,LP} & NDVI_{i,o} < NDVI_{i,LP} \end{cases}$$

donde  $NDVI_i^{Nj}$  corresponde al nuevo valor de NDVI para la fecha de captura  $i$  y número de iteración  $j$ .

- 5) Ajuste de la nueva serie temporal de NDVI mediante filtro SG ( $STAR^{Nj}$ ): se volvió a usar el filtro de SG sobre la nueva serie temporal generada en el punto anterior ( $STAi^{Nj}$ ). El objetivo en este caso era ajustar pequeñas variaciones en la nueva serie temporal (a diferencia del filtro aplicado en el punto 2, donde se pretendía obtener una serie temporal que mostrara la tendencia del cultivo a largo plazo). Para ello, los parámetros introducidos fueron diferentes: tamaño de ventana = 9 y un polinomio de grado 3 (obtenidos también a partir del CE).

- 6) Cálculo del índice de efecto de ajuste: este índice tiene como objetivo evaluar el grado en que los valores de NDVI ajustados ( $NDVI_{i,SG}^{Nj}$ ) se acercan a los puntos de NDVI de mayor ponderación. Dicho valor se calculó a partir de la siguiente fórmula:

$$F_j = \sum_{i=1}^n (|NDVI_{i,SG}^{Nj+1} - NDVI_{i,o}| \times W_i)$$

Hasta que el valor de  $F_j$  no cumpliera la condición de salida, se volvía a repetir el proceso de forma iterativa (desde el punto 4 al punto 6).

- 7) Condición de salida de iteración: Chen et al. (2004) sugirieron como condición de salida aquella iteración que minimizara  $F_j$ . Es decir, cuando se cumpliera la siguiente condición:

$$F_{j-1} \geq F_j \leq F_{j+1}$$

Para cuantificar la calidad de la técnica de relleno, volvió a calcular el RMSE entre los valores de NDVI coincidentes con las fechas en los que se rellenaron los huecos en la nueva serie temporal reconstruida (STAR<sub>S2</sub>) y la serie temporal interpolada (STi5) del CT.

Parte de los scripts generados para llevar acabo la metodología desarrollada en Python pueden consultarse en el ANEXO VII. SCRIPTS DE PYTHON.

## 5. RESULTADOS Y DISCUSIÓN

### 5.1. Selección del cálculo de índices a nivel pixel o a nivel parcela

La **Tabla 9** muestra la correlación media entre la opción de cálculo a nivel de pixel y a nivel de parcela de los diferentes cultivos para el cálculo de índices ópticos, mostrando en verde oscuro valores altos, y viceversa. Las correlaciones obtenidas son relativamente altas, superando el 0,99 en la mayoría de casos, llegando a ser de hasta 1 en el caso del índice SeLI. Los índices MCARI, MTCI y RDVI, muestran correlaciones ligeramente inferiores a los demás, bajando de 0,90 en el caso de MCARI. Estos valores menores de correlación van acompañados de valores altos de NRMSE (**Tabla 10**), pudiendo observarse en la mayoría de índices ópticos una relación inversa entre la correlación y el NRMSE medio. Sin embargo, no todos los índices cumplen dicha condición, como es el caso del CI-RE y NDMI que presentan valores altos tanto de correlación como de NRMSE. Esto indica que, para los índices NDVI, NDVIre, SAVI y SeLI, la forma en la que se calculen estos (a nivel de pixel o a nivel de objeto) no debería suponer diferencias significativas en los resultados (el NRMSE no llega a superar el 0,030), mientras que, calcular los índices CI-RE, MCARI, MTCI, NDMI RDVI y SLAVI a nivel de pixel o a nivel de objeto sí que podría implicar errores en los resultados.

En cuanto al análisis por cultivos, todos muestran una tendencia muy similar en cuanto a los valores de correlación media, pudiendo remarcar los cultivos de cebada, trigo y maíz como excepciones que presentan valores altos de NRMSE en el cálculo del índice MCARI.

**Tabla 9.** Correlación media entre la opción de cálculo a nivel de pixel y a nivel de parcela, para los 10 índices ópticos propuestos.

	CI-RE	MCARI	MTCI	NDMI	NDVI	NDVIre	RDVI	SAVI	SeLI	SLAVI
<b>CEBADA</b>	0,996	0,888	0,915	0,999	0,999	0,999	0,957	0,999	1	0,994
<b>TRIGO</b>	0,994	0,878	0,911	0,999	0,999	0,999	0,956	0,999	1	0,992
<b>GIRASOL</b>	0,997	0,926	0,923	1	0,999	0,999	0,956	0,999	1	0,996
<b>MAÍZ</b>	0,990	0,893	0,942	0,999	0,999	0,998	0,964	0,999	1	0,992
<b>VID</b>	0,996	0,898	0,874	0,999	0,999	0,998	0,941	0,999	1	0,994

**Tabla 10.** NRMSE medio entre la opción de cálculo a nivel de pixel y a nivel de parcela, para los 10 índices ópticos propuestos.

	CI-RE	MCARI	MTCI	NDMI	NDVI	NDVIre	RDVI	SAVI	SeLI	SLAVI
<b>CEBADA</b>	0,104	0,418	0,267	0,191	0,023	0,017	0,117	0,018	0,027	0,072
<b>TRIGO</b>	0,123	0,495	0,266	0,200	0,022	0,018	0,113	0,018	0,022	0,085
<b>GIRASOL</b>	0,087	0,273	0,273	0,125	0,019	0,017	0,131	0,016	0,020	0,081
<b>MAÍZ</b>	0,137	0,394	0,230	0,324	0,023	0,023	0,123	0,021	0,028	0,121
<b>VID</b>	0,059	0,210	0,216	0,196	0,019	0,014	0,100	0,014	0,019	0,025

La correlación media entre las dos opciones de cálculo para índices radar es ligeramente inferior que en los índices ópticos en la mayoría de los casos (**Tabla 11**), mostrando resultados muy similares entre ellos. En este caso hay una distinción un poco más clara entre los cultivos, siendo el trigo el que menor correlación muestra en el cálculo de índices IDPDD, RDIF, RMULTI y RSUM, llegando a correlaciones bajas de 0,734 (RMULTI). En cuanto a los valores medios de NRMSE (**Tabla 12**), la mayoría de los índices radar muestran valores superiores a los ópticos, llegando el índice IDPDD a obtener valores muy por encima de lo esperado (5,968 cuando lo normal es que los valores de NRMSE varíen de entre 0 y 1). El trigo es, de nuevo, el cultivo que vuelve a obtener valores más altos, coincidiendo el NRMSE máximo con el cálculo del índice RMULTI (índice que ha mostrado menor correlación). El cálculo de los índices RVI y VDDPI es el que presenta menor NRMSE, sin llegar a superar un error de 0,075 para ninguno de los cultivos.

Por lo tanto, podría decirse que realizar el cálculo del índice IDPDD a nivel de parcela implicaría cometer grandes errores (para los 5 cultivos), mientras que calculando el índice VDDPI con las dos opciones no deberían encontrarse grandes diferencias en los resultados (para los 5 cultivos). Las diferencias en el cálculo de los demás índices deberían mirarse más detalladamente para cada cultivo, intuyendo resultados similares en el cálculo de los índices RDIF y RSUM mediante las dos opciones en el girasol (además de VDDPI), siendo la vid el cultivo que mejores resultados presenta en conjunto.

**Tabla 11.** Correlación media entre la opción de cálculo a nivel de pixel y a nivel de parcela, para los 7 índices radar propuestos

	IDPDD	RDIF	RDIV	RMULTI	RSUM	RVI	VDDPI
<b>CEBADA</b>	0,979	0,979	0,980	0,986	0,987	0,985	0,980
<b>TRIGO</b>	0,899	0,899	0,974	0,734	0,908	0,980	0,974
<b>GIRASOL</b>	0,998	0,998	0,989	0,997	0,998	0,991	0,989
<b>MAÍZ</b>	0,940	0,940	0,913	0,913	0,950	0,966	0,913
<b>VID</b>	0,971	0,971	0,979	0,985	0,979	0,984	0,979

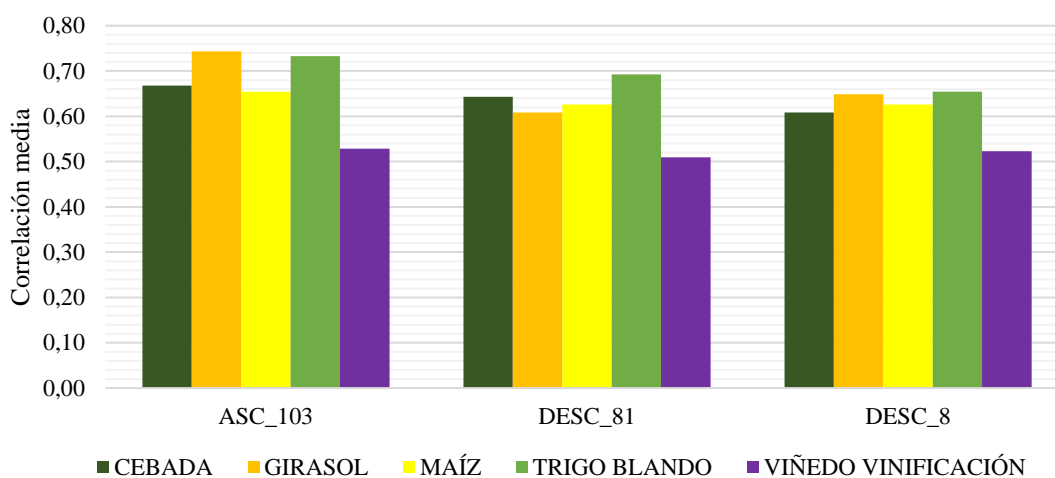
**Tabla 12.** NRMSE medio entre la opción de cálculo a nivel de pixel y a nivel de parcela, para los 7 índices radar propuestos.

	IDPDD	RDIF	RDIV	RMULTI	RSUM	RVI	VDDPI
<b>CEBADA</b>	5,515	0,071	0,106	0,142	0,052	0,062	0,017
<b>TRIGO</b>	5,503	0,247	0,112	0,789	0,194	0,068	0,018
<b>GIRASOL</b>	5,968	0,014	0,094	0,051	0,011	0,057	0,013
<b>MAÍZ</b>	5,700	0,140	0,145	0,268	0,105	0,073	0,023
<b>VID</b>	5,782	0,055	0,087	0,088	0,041	0,053	0,013



## 5.2. Selección de pasada de Sentinel-1

La **Figura 13** muestra la correlación media de las parcelas por cultivo entre los índices de Sentinel-1 y Sentinel-2 para las tres pasadas de Sentinel-1. La pasada ascendente (ASC\_103) es la que mayor correlación presenta entre ambos índices para los cinco cultivos. Sin embargo, el número reducido de capturas que presenta a lo largo del período de estudio comparado con las otras dos pasadas (**Figura 6**) hizo que se descartara. Entre las dos pasadas descendentes, la DESC\_81 presentaba mayor correlación para los cultivos de la cebada y el trigo, y menor para los del girasol y la vid (el maíz presentó correlaciones iguales para ambas pasadas). Teniendo en cuenta que la pasada DESC\_81 contaba con un mayor número de pasadas en el periodo de estudio (**Figura 6**), se decidió seleccionar esta última para realizar el trabajo.



**Figura 13.** Correlación media de parcelas por cultivo entre los índices de Sentinel-1 y Sentinel-2 para las tres pasadas de Sentinel-1.

## 5.3. Selección de índices

La **Figura 14** muestra una matriz de correlaciones medias de todas las parcelas de cebada entre los índices de Sentinel-1 y Sentinel-2. Empezando a analizar los índices de Sentinel-1, puede observarse que los índices IDPDD, RDIF, RSUM y VV muestran correlaciones muy similares con los índices de Sentinel-2 muy. Otro grupo de índices con comportamiento similar sería el RDIV, RVI y VDDPI. Los índices RMULTI y VH son los únicos que no tienen un comportamiento semejante a los demás. En definitiva, se pueden realizar agrupaciones con los índices que muestran correlaciones altas entre sí (**Tabla 13**).

**Tabla 13.** Grupos de índices de Sentinel-1 con comportamiento de correlación similar frente a índices de Sentinel-2.

Nombre grupo	Índices pertenecientes al grupo
RDIF	IDPDD, RDIF, RSUM, VV
RDIV	RDIV, RVI, VDDPI
RMULTI	RMULTI
VH	VH

IDPDD	0.55	0.38	0.56	0.49	0.42	0.68	0.55	0.41	0.51	0.52
RDIF	-0.55	-0.38	-0.56	-0.48	-0.42	-0.68	-0.55	-0.41	-0.51	-0.52
RDIV	0.59	0.52	0.59	0.65	0.52	0.76	0.6	0.51	0.55	0.61
RMULTI	-0.4	-0.26	-0.4	-0.28	-0.31	-0.41	-0.39	-0.31	-0.38	-0.35
RSUM	-0.51	-0.33	-0.51	-0.41	-0.39	-0.6	-0.5	-0.38	-0.47	-0.47
RVI	0.59	0.51	0.59	0.65	0.51	0.76	0.59	0.5	0.55	0.61
VDDPI	0.59	0.52	0.59	0.65	0.52	0.76	0.6	0.51	0.55	0.61
VH	-0.19	-0.064	-0.17	-0.019	-0.12	-0.1	-0.14	-0.11	-0.16	-0.13
VV	-0.53	-0.36	-0.54	-0.45	-0.41	-0.64	-0.53	-0.4	-0.49	-0.49
	CI-RE	MCARI	MTCI	NDMI	NDVI	NDVIre	RDVI	SAVI	SeLI	SLAVI

**Figura 14.** Matriz de correlaciones medias de las parcelas del cultivo de cebada entre los índices Sentinel-1 (pasada DESC\_81) y Sentinel-2.

En cuanto a los índices de Sentinel-2, puede verse como los índices CI-RE, MTCI y RDVI muestran correlaciones similares frente a los índices de Sentinel-1 entre sí, así como los índices NDMI, SAVI, SeLI y SLAVI. Los índices NDVI y SAVI muestran también correlaciones similares entre sí, quedando los índices NDVIre y MCARI sin correlaciones significativas con los demás índices. Los grupos finales de índices de Sentinel-2 quedarían, por lo tanto, tal y como muestra la **Tabla 14**.

**Tabla 14.** Grupos de índices de Sentinel-2 con comportamiento de correlación similar frente a índices de Sentinel-1.

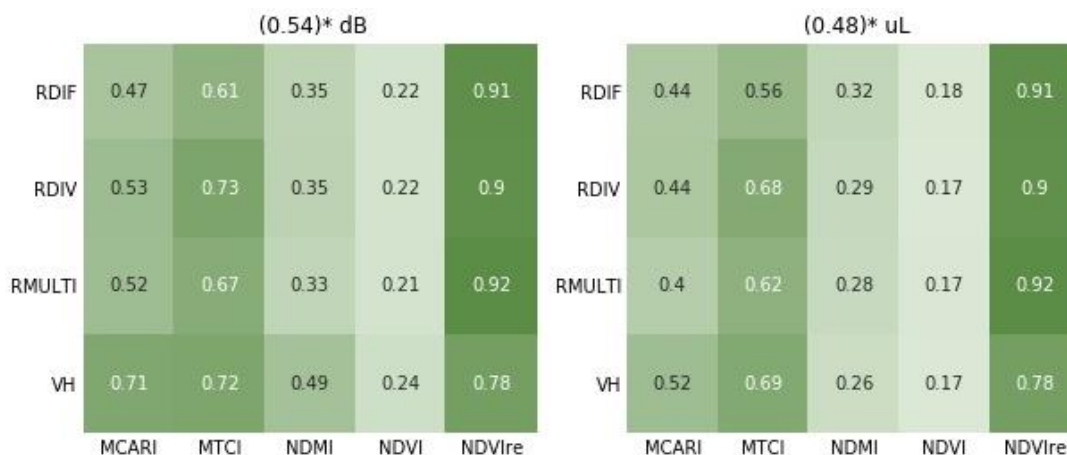
Nombre grupo	Índices pertenecientes al grupo
MTCI	MTCI, CI-RE, RDVI
NDMI	NDMI, SeLI, SLAVI
NDVI	NDVI, SAVI
NDVIre	NDVIre
MCARI	MCARI

Aunque las correlaciones medias variaran entre cultivos, los comportamientos analizados en la matriz de correlaciones medias de la **Figura 14** también se observaron en las matrices de correlaciones medias de los cultivos de interés, por lo que se decidió trabajar con estas agrupaciones para todos los cultivos (véase ANEXO II. MATRICES DE CORRELACIÓN para consultar las matrices de correlaciones medias de los cultivos de interés).

#### 5.4. Selección de unidades

La **Figura 15** muestra dos matrices de correlaciones medias (en valor absoluto), entre los índices de Sentinel-1 en unidades de decibelios (dB, izquierda) y lineales ( $m^2m^{-2}$ , derecha), y Sentinel-2 correspondientes a las 75 parcelas de cebada en regadío del conjunto de entrenamiento. Puede observarse como la correlación entre índices es mayor cuando los índices radar se calculan en dB, a excepción del índice óptico NDVIre, que muestra los mismos coeficientes de correlación en

ambos casos. Este índice es, además, el que presenta mayores correlaciones con los índices radar, obteniendo la correlación máxima con el índice RMULTI (0,92). Los números en paréntesis situados en la parte superior de cada matriz de correlación indican la correlación media de la matriz, siendo superior al utilizar unidades de dB (0,54) frente a  $m^2m^{-2}$  (0,48). Sin embargo, aunque utilizar unidades de dB presente mejores resultados, las correlaciones máximas entre índices (en este caso, tanto para unidades de dB como  $m^2m^{-2}$ , obtenidas con el índice NDVIre) se mantienen en las dos opciones.



**Figura 15.** Matrices de correlaciones medias (en valor absoluto) de las parcelas del CE de la cebada en regadío entre los índices de Sentinel-1 y Sentinel-2 en unidades de decibelios (dB, izquierda) y lineales ( $m^2m^{-2}$ , derecha). \*Media de la matriz de correlaciones medias.

La cebada en secano, el trigo en secano y en regadío y la vid muestran también medias de correlación mayores cuando el cálculo de índices de Sentinel-1 se realiza en unidades de dB (0,52 en dB vs 0,47 en  $m^2m^{-2}$ ; 0,67 vs. 0,62; 0,56 vs 0,55 y 0,38 vs 0,37 respectivamente), mientras que el girasol mejora sus resultados en  $m^2m^{-2}$  (0,48 vs 0,51). El maíz muestra medias iguales para las dos opciones de cálculo diferente (0,51). Para los cultivos de invierno y la vid, la correlación media más alta se obtiene entre los índices RMULTI y NDVIre, mientras que, para los cultivos de verano, la mejor correlación se alcanza entre los índices RDIV y NDVIre. La diferencia máxima entre las correlaciones medias de las dos opciones de cálculo es muy pequeña (0,05), Además, puede observarse en la **Figura 15** como, a pesar de ser mayor la media en dB, los máximos entre índices se mantienen. Todo ello indica que, para los cultivos con los que se está trabajando, las unidades en las que se realice el cálculo de los índices de Sentinel-1 no debería suponer una mejora o un empeoramiento significativo en los resultados. La **Tabla 15** muestra un resumen con los datos comentados de todos cultivos (véase ANEXO II. MATRICES DE CORRELACIÓN para consultar las matrices de correlaciones medias del resto de cultivos):

**Tabla 15.** Datos de interés de las matrices de correlaciones medias de comparación de unidades (dB vs.  $m^2m^{-2}$ ) para cada cultivo, indicando las unidades con mayor media de correlaciones y la mejor combinación de índices.

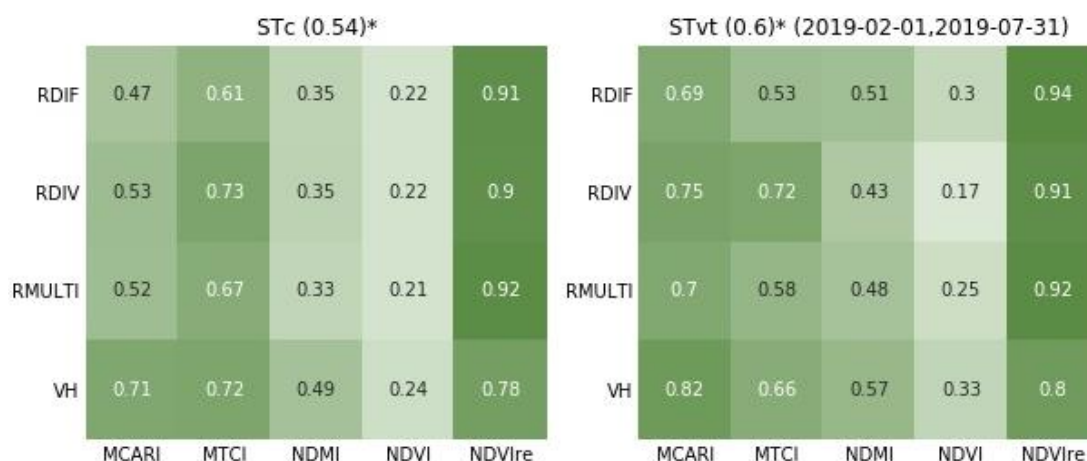
CULTIVO	C. MEDIA MATRIZ*		COMBINACIÓN ÓPTIMA		
	dB	$m^2m^{-2}$	Índices	C. Max.**	Unidades
CEBADA REGADÍO	0,54*	0,48	RMULTI-NDVIre	0,92	dB, $m^2m^{-2}$
CEBADA SECANO	0,52	0,47	RDIF/RMULTI-NDVIre	0,95	
GIRASOL	0,48	0,51	RDIV-NDVIre	0,92	
MAÍZ		0,51	RDIV-NDIVre	0,92	
TRIGO REGADÍO	0,56	0,55	RMULTI-NDVIre	0,87	
TRIGO SECANO	0,67	0,62	RMULTI-NDVIre	0,96	
VIÑEDO VINIFICACIÓN	0,38	0,37	RDIV/RMULTI-NDVIre	0,86	

\*: correlación media de la matriz. \*\*: correlación media máxima entre índices.

En el presente estudio se decidió trabajar en unidades de dB.

### 5.5. Selección de ventanas temporales

La **Figura 16** muestra dos matrices de correlaciones medias (en valor absoluto) entre los índices de Sentinel-1 y Sentinel-2: la matriz de la izquierda muestra la correlación para toda la serie temporal, mientras que la matriz de la derecha muestra la correlación para una ventana temporal ( $ST_{VT}$ ) ajustada al ciclo fenológico del cultivo de la cebada en regadío. Puede verse que la correlación aumenta cuando la serie temporal se ajusta al calendario del cultivo para todos los índices, siendo de nuevo el índice NDVIre el que mayor correlación presenta con los índices radar, obteniendo, en esta ocasión, la correlación máxima con el índice radar RDIF (0,94).



**Figura 16.** Matrices de correlaciones medias (en valor absoluto) de las parcelas del CE de la cebada en regadío entre los índices de Sentinel-1 y Sentinel-2, con datos de toda la serie temporal ( $ST_C$ , izquierda) y calendario fenológico de la cebada ( $ST_{VT}$ , derecha). \*Media de la matriz de correlaciones medias.

La **Tabla 16** muestra un breve resumen con los datos más importantes extraídos de las matrices de correlaciones medias de los cultivos (tanto el girasol como el trigo en regadío siguen el mismo comportamiento, en este caso, que la cebada en regadío). Cabe remarcar que, a diferencia de los demás cultivos, las medias de las correlaciones de la cebada en secano y del trigo en secano no aumentan al aplicar la ventana temporal. Esto podría deberse a diferentes razones, entre ellas, la

influencia positiva que pudieran tener valores altos de correlaciones correspondientes a cultivos precedentes o posteriores en la misma parcela, careciendo de interés para el presente estudio. Es por ello que se decide trabajar con la serie temporal ajustada a cada ciclo fenológico de cada cultivo. De nuevo, la combinación predominante que presenta valores de correlaciones mayores es RMULTI-NDVI<sub>Ire</sub>, alcanzando valores altos también las combinaciones RDIV-NDVI<sub>Ire</sub> y RDIF-NDVI<sub>Ire</sub>.

**Tabla 16.** Datos de interés de las matrices de correlaciones medias de comparación de ventanas temporales (ST<sub>C</sub> vs. ST<sub>VT</sub>) para cada cultivo, indicando la mejor opción y la mejor combinación de índices.

CULTIVO	C. MEDIA MATRIZ*		COMBINACIÓN ÓPTIMA		
	ST <sub>C</sub>	ST <sub>VT</sub>	Índices	C. Max.**	VT
CEBADA REGADÍO	0,54	0,60	RDIF-NDVI <sub>Ire</sub>	0,94	ST <sub>VT</sub>
CEBADA SECANO	0,52	0,51	RDIF/RMULTI-NDVI <sub>Ire</sub>	0,95	ST <sub>C</sub>
GIRASOL	0,51	0,70	RDIV/RMULTI-NDVI <sub>Ire</sub>	0,95	ST <sub>VT</sub>
MAÍZ	0,51	0,48	RDIV-NDIV <sub>Ire</sub>	0,95	ST <sub>VT</sub>
TRIGO REGADÍO	0,56	0,59	RMULTI-NDVI <sub>Ire</sub>	0,88	ST <sub>VT</sub>
TRIGO SECANO	0,67	0,58	RMULTI-NDVI <sub>Ire</sub>	0,96	ST <sub>C</sub>

\*: correlación media de la matriz. \*\*: correlación media máxima entre índices.

Al ser la vid un cultivo permanente no se aplicó ninguna ventana temporal específica, sino que se continuó trabajando con toda la serie temporal disponible.

## 5.6. Selección de suavizado de curvas

La **Figura 17** muestra las matrices de correlaciones medias entre los índices Sentinel-1 y Sentinel-2 para la serie temporal suavizada mediante el filtro SG con una ventana temporal de 15 días (izquierda) y 31 días (derecha), para el cultivo de la cebada en regadío. Como era de esperar, cuanto mayor es el tamaño de ventana del filtro SG, mayor correlación existirá entre los índices de S1 y S2 debido a esa eliminación del *ruido* (o variabilidad a corto plazo) ya comentada. Puede verse cómo la media de la matriz de correlaciones aumenta hasta un 0,70 con una ventana temporal de 31 días, siendo de nuevo el índice óptico NDVI<sub>Ire</sub> el que mayores correlaciones alcanza, llegando a un 0,97 con el índice radar RDIF.

En este caso, tal y como se muestra en la **Tabla 17**, la media de la matriz mejora para todos los cultivos al calcular las correlaciones entre los índices de S1 y S2 con el filtro SG aplicado. Puede verse que las medias de las matrices de correlaciones aumentan con el tamaño de ventana de 31 días, así como el valor de correlación máximo obtenido con la mejor combinación de índices S1 y S2. Una vez más las correlaciones más altas vuelven a obtenerse con la combinación del índice NDVI<sub>Ire</sub> de Sentinel-2, combinado con el índice radar RMULTI para el cultivo de cebada en seco (mismo valor con RDIF), girasol (mismo valor con RDIV), maíz y trigo en regadío. Para el caso de la cebada en regadío y el trigo en seco, la mayor correlación la obtiene junto con el índice RDIF, y, la vid, con VH.



**Figura 17.** Matrices de correlaciones medias de las 75 parcelas del CE de la cebada en regadío entre los índices de Sentinel-1 y Sentinel-2, con filtro SG aplicado para diferente tamaño de ventana: 15 ( $ST_{VT,15}$ , izquierda) y 31 ( $ST_{VT,31}$ , derecha). \*Media de la matriz de correlaciones medias.

Llama la atención como para los cultivos de invierno, las medias de las matrices obtenidas sean mayores en secano que en regadío, aunque, profundizando un poco más, pueda observarse como, en el caso de la cebada, por ejemplo, la diferencia entre los coeficientes medios de correlación máximos obtenidos es menor (0,97 en regadío y 0,95 en secano) que la media de las matrices (0,70 en regadío y 0,57 en secano). Sin embargo, en el caso del trigo, aunque en regadío se obtenga una media de la matriz mayor, se obtiene un coeficiente medio de correlación máximo mayor en secano (0,93 en regadío y 0,96 en secano). A modo resumen, son la cebada en regadío, girasol y maíz los cultivos que mayor coeficiente de correlación presentaron (0,97), seguido del trigo en secano (0,96), la cebada en secano (0,95), el trigo en regadío (0,93) y la vid (0,89). Los altos coeficientes de correlación que se han obtenido indican, a priori, que la reconstrucción de la serie temporal de índices ópticos a partir de datos de índices radar parece ser factible.

**Tabla 17.** Datos de interés de las matrices de correlaciones medias de comparación de tamaño de ventanas aplicadas mediante filtro SG: sin filtro, 15 y 31 días ( $ST_{VT}$  vs.  $ST_{VT,15}$  vs.  $ST_{VT,31}$ ) para cada cultivo, indicando la mejor opción y la mejor combinación de índices.

CULTIVO	C. MEDIA MATRIZ*			COMBINACIÓN ÓPTIMA		
	$ST_{VT}$	$ST_{VT,15}$	$ST_{VT,31}$	Índices	C. Max.**	Suavizado
CEBADA REGADÍO	0,60	0,64	0,70	RDIF-NDVIre	0,97	
CEBADA SECANO	0,51	0,54	0,57	RDIF/RMULTI-NDVIre	0,95	
GIRASOL	0,70	0,75	0,80	RDIV/RMULTI-NDVIre	0,97	
MAÍZ	0,48	0,54	0,59	RMULTI-NDIVre	0,97	$ST_{VT,31}$
TRIGO REGADÍO	0,59	0,65	0,72	RMULTI-NDVIre	0,88	
TRIGO SECANO	0,58	0,61	0,65	RDIF-NDVIre	0,96	
VIÑEDO VINIFICACIÓN	0,38	0,63	0,72	VH-NDVIre	0,89	

\*: correlación media de la matriz. \*\*: correlación media máxima entre índices.



### 5.7. Análisis del comportamiento reflexivo y retrodispersivo de los cultivos

A partir del suavizado de curvas (de índices ópticos y radar) realizado mediante el filtro SG, se obtienen la mediana e IQR de las series temporales suavizadas (**Figura 18**) con una ventana de 31 días y normalizadas, para una ventana temporal comprendida entre 01/02/2019 y 31/07/2019 para la cebada en regadío, mostrando la correlación entre las diferentes combinaciones. Las zonas sombreadas verticales muestran diferentes fases de desarrollo del cultivo (basados en la **Tabla 3**). Correspondiendo el color verde oscuro al periodo vegetativo, el rosa claro al periodo reproductivo, rosa oscuro al llenado y madurez del fruto y el amarillo a la cosecha. El comportamiento de la cebada en regadío es muy parecido para todos los índices de Sentinel-2, coincidiendo un aumento de sus valores en el periodo vegetativo, alcanzando los valores máximos en mitad de esta fase (coincidiendo con el final de la fase de crecimiento del tallo). A partir de este punto, con el comienzo del crecimiento de la inflorescencia (todavía en la fase reproductiva), el valor de los índices de S2 comienzan a decaer, hasta alcanzar los mínimos en la época de cosecha.

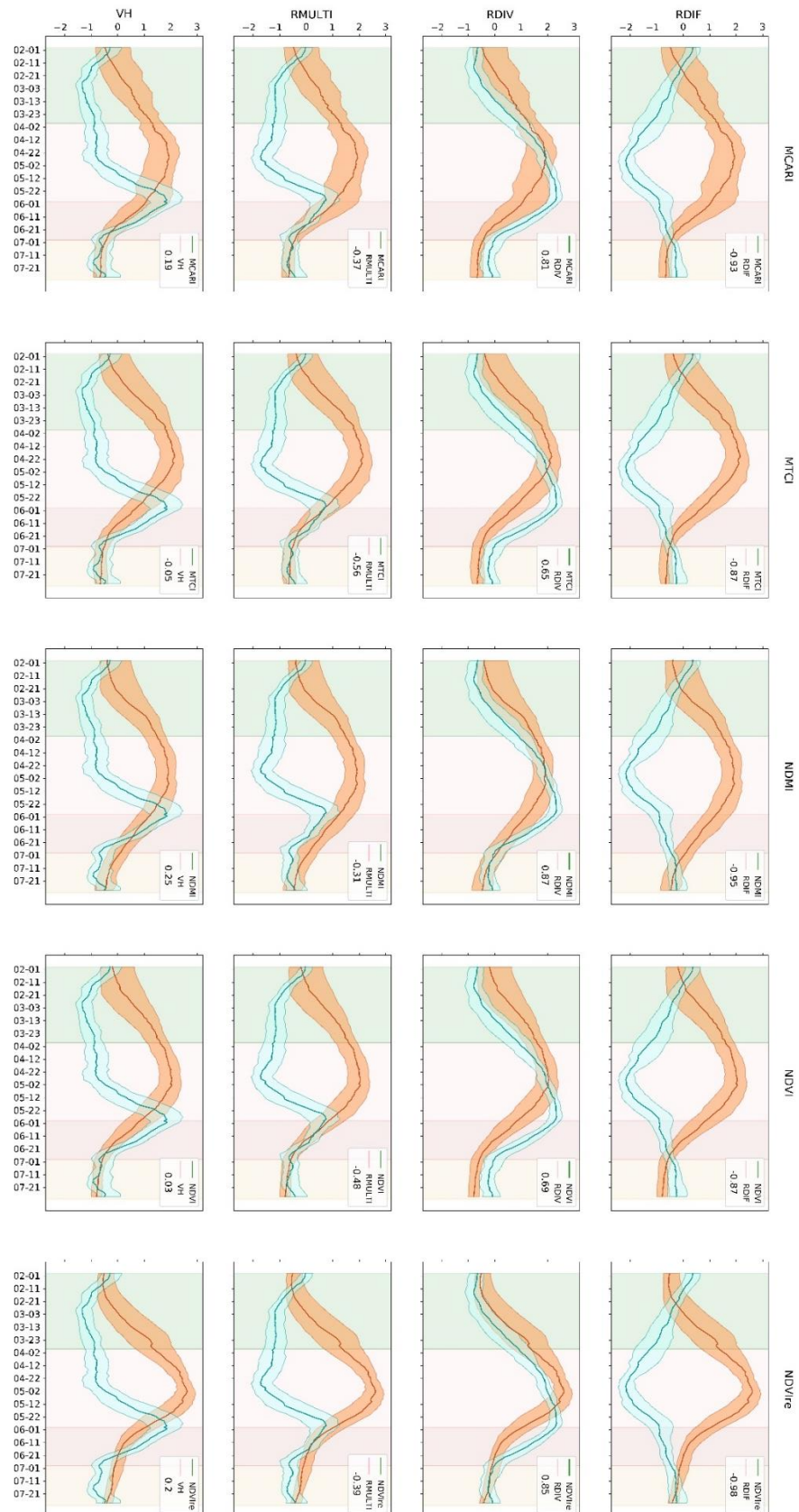
Entre los índices de S2, el NDVI<sub>re</sub> es el que mayor sensibilidad muestra a los cambios en el ciclo del cultivo, generando variaciones en la curva más acentuadas, obteniendo, además, correlaciones mayores con los índices radar. En cuanto a estos últimos se observa que, a diferencia de los índices ópticos, no todos muestran comportamientos similares para el mismo cultivo. Por un lado, los índices VH y RMULTI describen un comportamiento bastante similar: no se produce gran depolarización a medida que crece el cultivo, hasta el pico en mayo (coincidiendo con el fin del crecimiento vegetativo y el inicio de la floración y fructificación). En ese momento, la espiga de la cebada aumenta tanto su tamaño como su peso, produciendo el conocido *tumbado* de la cebada, perdiendo así la geometría vertical de esta, y, por lo tanto, aumentando la retrodispersión de volumen representada por estos índices. Por otro lado, el índice RDIV muestra cierta sensibilidad ante las variaciones de cultivo, trazando un patrón semejante a las curvas de los índices de S2, aunque con cierto retraso temporal, obteniendo el máximo a final de la fase de crecimiento de la inflorescencia y no a su comienzo como los índices de S2. En cuanto al índice RDIF, sigue una tendencia inversa a los índices de S2 y a RDIV, alcanzando su mínimo al final de la fase de crecimiento del tallo. En esta fase de encañado, la retrodispersión en VV<sup>1</sup> baja debido a la verticalidad de los tallos, ya que estos absorben en mayor medida el pulso radar cuando la polarización es vertical. Se obtiene que la combinación de índices que muestra mayor correlación media es la RDIF-NDVI<sub>re</sub>, con un valor de -0,97. La cebada en secano presenta comportamientos muy similares, mostrando cierto adelantamiento en los picos máximos de las curvas (ver ANEXO

---

<sup>1</sup> De media, el valor de retrodispersión directa suele ser hasta diez veces mayor que el valor de retrodispersión cruzada, pudiendo considerar, por lo tanto, que el índice RDIF refleja el comportamiento retrodispersivo de VV.

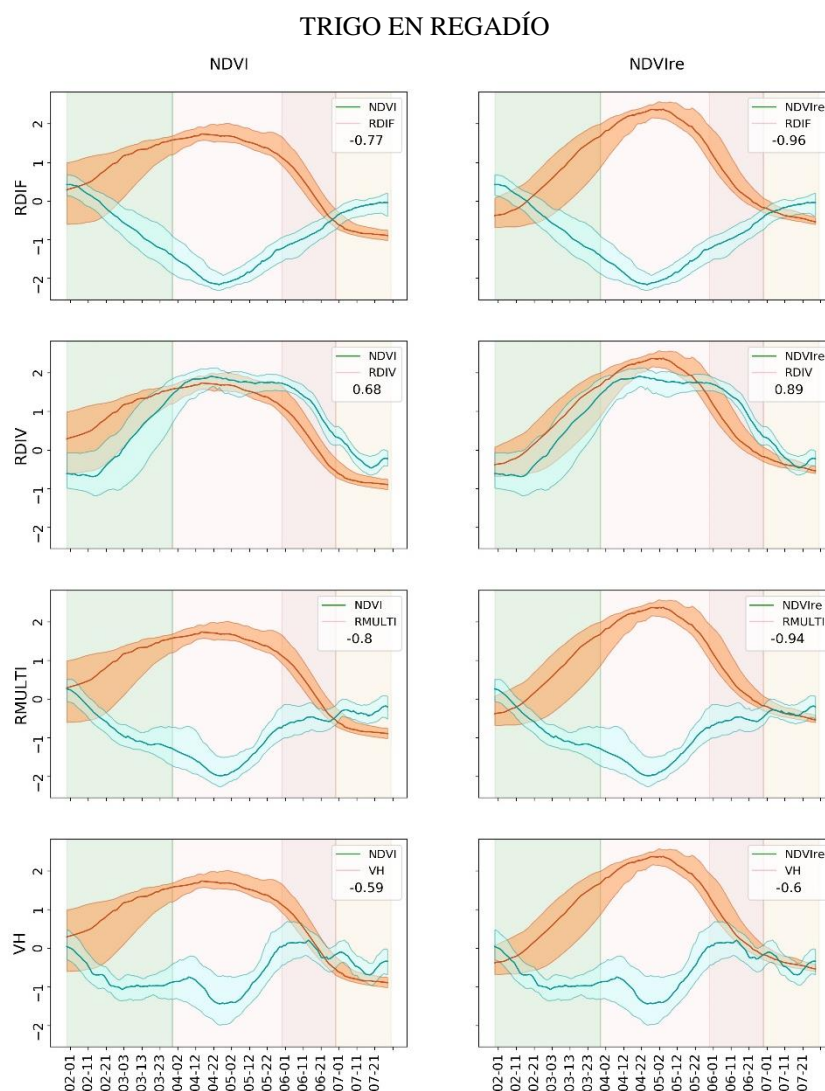


### III. COMPORTAMIENTO REFLEXIVO Y RETRODISPERSIVO DE LOS CULTIVOS para consultar la figura correspondiente).



**Figura 18.** Mediana e IQR (sombreado) de las series temporales (01/02/2019 – 31/07/2019) suavizadas mediante el filtro SG (ventana de 31 días) y normalizadas, para los índices de Sentinel-1 (azul) y Sentinel-2 (naranja), para el cultivo de cebada en regadío. Zonas sombreadas: período vegetativo, período reproductivo, llenado y madurez, cosecha.

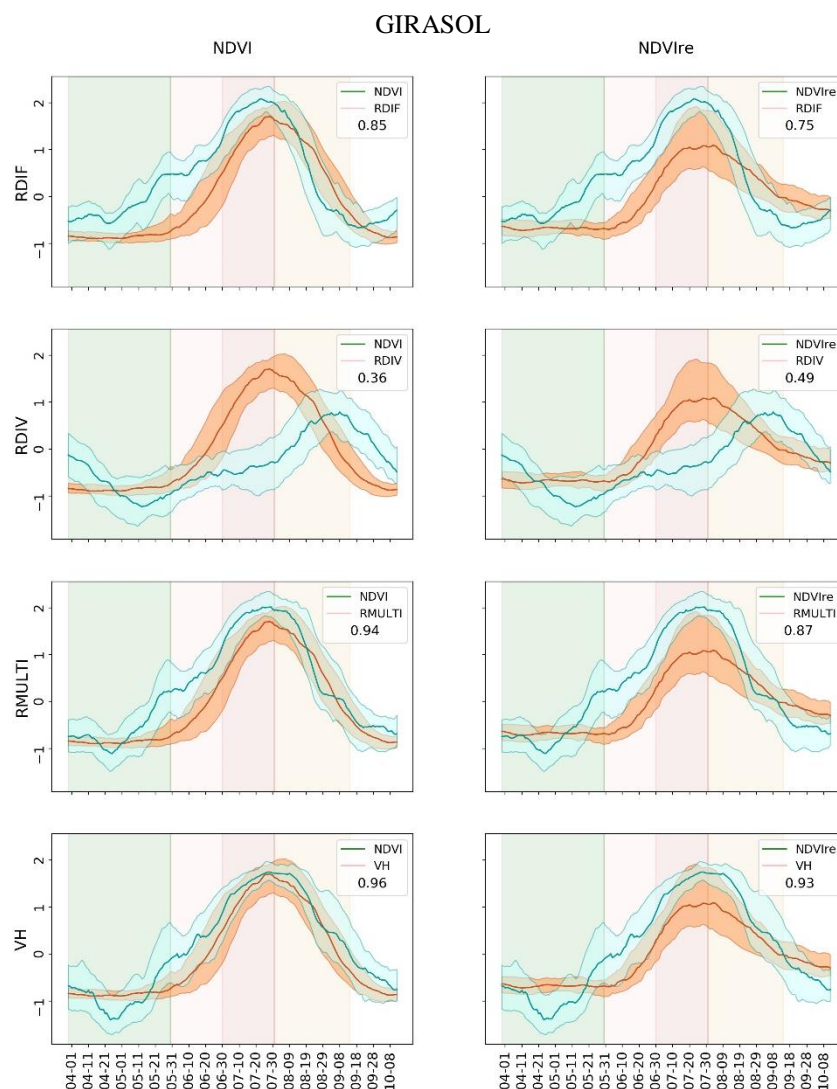
Puede observarse que, tanto en la **Figura 18** y las siguientes figuras que muestran la mediana e IQR de las series temporales suavizadas mediante el filtro SG, las correlaciones entre diferentes combinaciones de índices que muestran no son las mismas que las obtenidas en las matrices de correlación del apartado 5.6. Esto se debe al modo en el que se ha realizado el cálculo de las correlaciones: en el apartado 5.6 se han calculado las correlaciones para todas las parcelas de forma individual y, posteriormente, se realiza la media. En el segundo caso, se obtiene la serie temporal mediana de los índices (para todas las parcelas) y, a continuación, se calcula la correlación entre ellos.



**Figura 19.** Mediana e IQR (sombreado) de las series temporales (01/02/2019 – 31/12/2019) suavizadas mediante el filtro SG (ventana de 31 días) y normalizadas, para los índices de Sentinel-1 (azul) y Sentinel-2 (naranja), para el cultivo de trigo en regadío. Zonas sombreadas: período vegetativo, período reproductivo, llenado y madurez, cosecha.

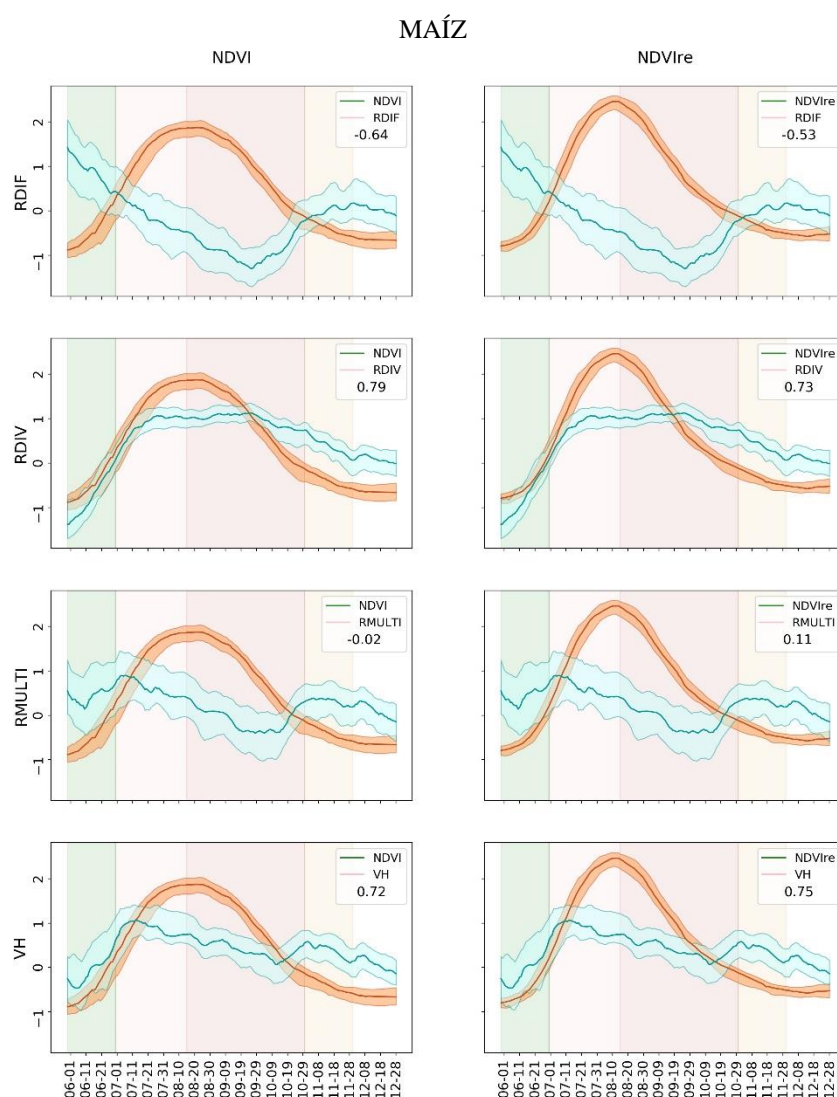
La **Figura 19** muestra las curvas suavizadas para el trigo. En este caso, y para los cultivos restantes, se representan únicamente las relaciones entre los índices ópticos NDVI (debido a su

amplio uso en el campo agronómico) y NDVIre (por presentar altas correlaciones) con los 4 índices radar con el fin de acotar los resultados (ver ANEXO III. COMPORTAMIENTO REFLEXIVO Y RETRODISPERSIVO DE LOS CULTIVOS para consultar los demás índices ópticos). En este caso puede verse como el trigo sigue un comportamiento muy parecido a la cebada. Las diferencias se observan en los índices RMULTI y VH, ya que en el trigo no ocurre el *tumbado* que se ha comentado. Los picos de valores de mínimos en RDIF representan la verticalidad de los tallos del trigo, ya que estos absorben en mayor medida el pulso radar en polarización VV. Además, cuanto mayor sea la biomasa del tallo, mayor será la absorción, coincidiendo, por lo tanto, el pico mínimo de RDIF con el máximo de NDVI. La correlación máxima se vuelve a alcanzar con la combinación de los índices RDIF-NDVIre, con un valor de correlación de -0,96.



**Figura 20.** Mediana e IQR (sombreado) de las series temporales (01/04/2019 – 15/10/2019) suavizadas mediante el filtro SG (ventana de 31 días) y normalizadas, para los índices de Sentinel-1 (azul) y Sentinel-2 (naranja), para el cultivo de girasol. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

A diferencia de los cultivos de invierno, el girasol (**Figura 20**) presenta una relación directa entre todos los índices de S1 y S2, mostrando los índices RDIF, RMULTI y VH tendencias muy similares, y, por lo tanto, correlaciones también semejantes, alcanzando los picos máximos en el final del periodo de llenado y madurez y comienzo de cosecha. La mayor correlación entre índices de S1 y S2 se alcanza con la combinación VH-NDVI, con un valor de 0,96. El índice RDIV es el que menor correlación presenta con los índices de S2, siguiendo una tendencia plana, sin producirse grandes depolarizaciones hasta el máximo que se observa en el periodo de cosecha.

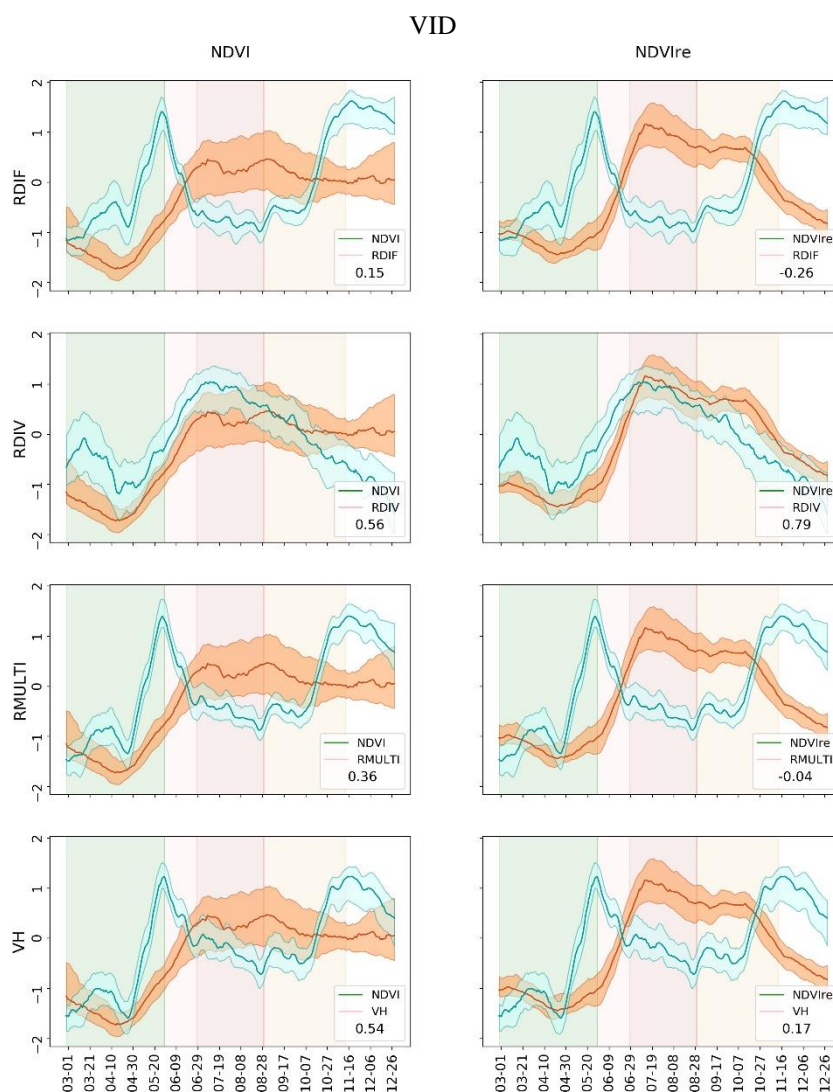


**Figura 21.** Mediana e IQR (sombreado) de las series temporales (01/07/2019 – 31/12/2019) suavizadas mediante el filtro SG (ventana de 31 días) y normalizadas, para los índices de Sentinel-1 (azul) y Sentinel-2 (naranja), para el cultivo de maíz. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

El comportamiento de los índices ópticos en la vid (**Figura 22**) difieren un poco más de lo que se ha ido viendo hasta el momento, siendo el NDVIre el que presenta mayor sensibilidad frente a los cambios en la vegetación, mostrando picos máximos al comienzo del periodo de llenado y



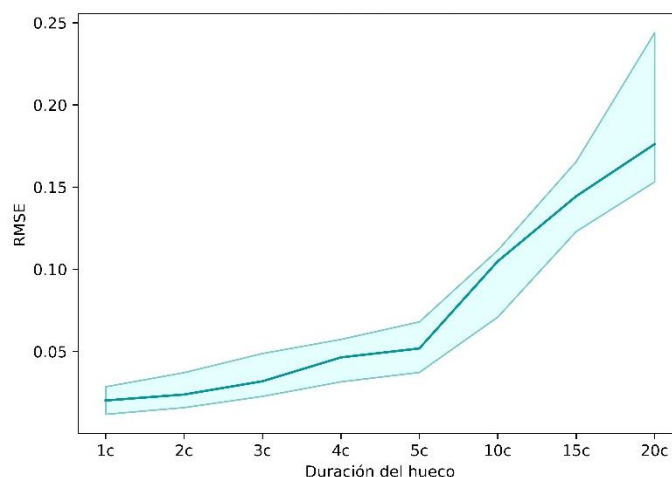
madurez, seguido de una tendencia descendente hasta el periodo de cosecha. El NDVI, sin embargo, refleja dos picos máximos (uno al comienzo y otro al fin del periodo reproductivo), manteniéndose constante en el periodo de cosecha, tras sufrir un ligero descenso después del segundo pico mencionado. Los índices radar muestran un comportamiento curioso, ya que el RDIF, RMULTI y VH reflejan un primer ascenso pronunciado en el periodo vegetativo, alcanzando su máximo al comienzo del periodo reproductivo, seguido de un descenso rápido que alcanza valores constantes hasta el comienzo de la cosecha, donde vuelven a verse incrementados los valores. En cuanto al índice RDIV, puede observarse como sigue de forma más clara las dinámicas de los índices ópticos.



**Figura 22.** Mediana e IQR (sombreado) de las series temporales (01/03/2019 – 31/12/2019) suavizadas mediante el filtro SG (ventana de 31 días) y normalizadas, para los índices de Sentinel-1 (azul) y Sentinel-2 (naranja), para el cultivo de la vid. Zonas sombreadas: período vegetativo, período reproductivo, llenado y madurez, cosecha.

## 5.8. Evaluación del impacto de los huecos

La **Figura 23** muestra la evolución del RMSE de las parcelas de test (CT), en función de la duración de los huecos para el NDVI de la cebada en regadío. Se observa que a medida que la duración del hueco es mayor el RMSE aumenta consistentemente. Esta tendencia ascendente también se obtiene con los demás cultivos, aunque los valores varíen, tal y como se resume en la **Tabla 18**, donde los verdes más oscuros presentan valores altos de RMSE y viceversa.



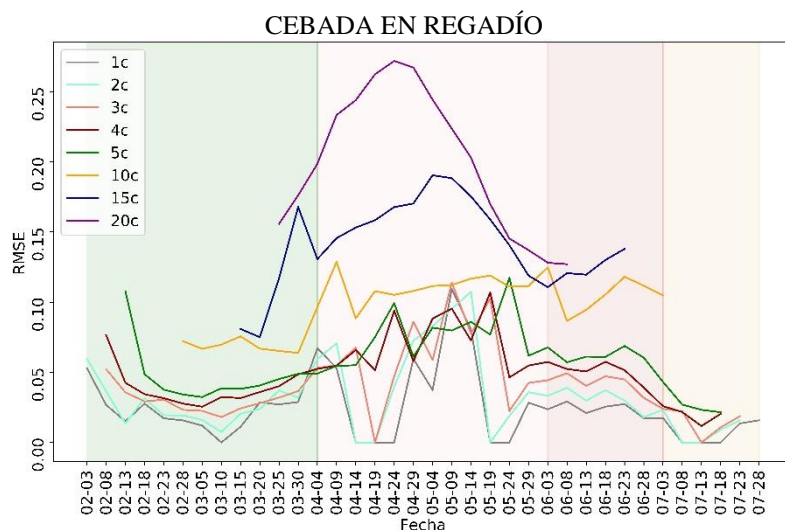
**Figura 23.** Representación de la mediana e IQR (sombreado) del RMSE para las parcelas del CT de cebada en regadío en función de la duración de los huecos: 1c (un captura sin información), 2c (dos capturas seguidas sin información), etc. hasta 20c, a lo largo de las ventanas temporales definidas.

La vid es el cultivo que presenta menor RMSE, con un valor máximo de 0,038 alcanzado cuando la duración máxima de los huecos (20c, 20 capturas seguidas sin información), seguido del trigo en secano (RMSE máximo de 0,112). Las diferencias de RMSE obtenidos entre secano y regadío no son muy significativas, obteniendo una diferencia máxima de 0,042 en el trigo. Los cultivos de verano (girasol y maíz), sin embargo, sí que siguen una tendencia de aumento de valores similares, con diferencias máximas de hasta 0,026 en huecos de duración máxima. Otro detalle interesante a remarcar es el notable aumento en el RMSE en el salto de 5c a 10c, con excepción de la vid, que se mantiene bastante constante. Tanto los datos presentados en la **Tabla 18** como las figuras incluidas en el ANEXO IV. RMSE DE EVALUACIÓN DE IMPACTO DE HUECOS indican una clara tendencia de aumento del RMSE a medida que la duración de los huecos se alarga.

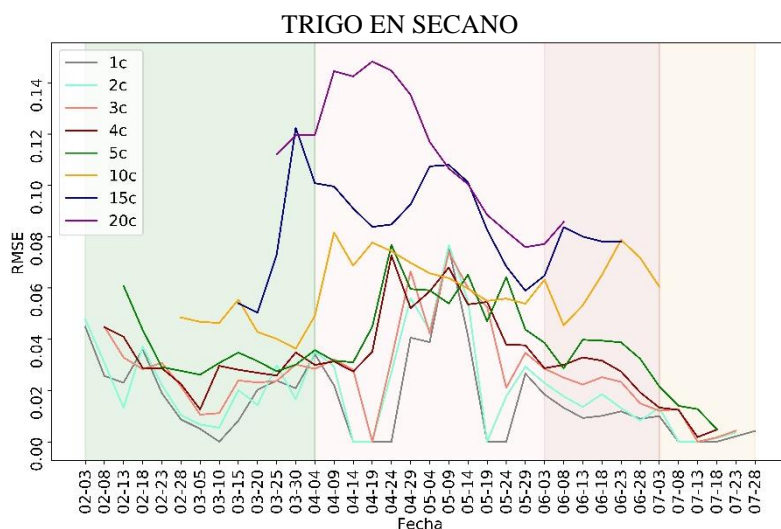
**Tabla 18.** Resumen de las medianas de RMSE en NDVI para las parcelas de CT de los cultivos de interés, en la ventana temporal definida para cada cultivo.

RMSE	1c	2c	3c	4c	5c	10c	15c	20c
CEBADA REGADÍO	0,020	0,024	0,032	0,046	0,052	0,105	0,144	0,176
CEBADA SECANO	0,013	0,014	0,021	0,027	0,031	0,062	0,103	0,140
TRIGO B. REGADÍO	0,021	0,024	0,033	0,043	0,052	0,091	0,133	0,164
TRIGO B. SECANO	0,013	0,017	0,024	0,029	0,032	0,055	0,084	0,112
GIRASOL	0,021	0,027	0,039	0,049	0,059	0,130	0,185	0,235
MAÍZ	0,019	0,027	0,041	0,057	0,072	0,105	0,164	0,209
VIÑEDO VINIFICACIÓN	0,009	0,012	0,016	0,018	0,020	0,027	0,033	0,038

Las **Figura 24** y **Figura 25** representan el RMSE en NDVI en función del momento en el que ocurren los huecos, para la cebada en regadío (comportamiento similar y por lo tanto representativo de la cebada en secano) y el trigo en secano (comportamiento similar y por lo tanto representativo del trigo en regadío). La fecha del hueco de las figuras corresponde a la fecha del punto medio del hueco, es decir, el RMSE de un hueco de duración de, por ejemplo, 5 capturas (5c), entre los días 18 de febrero y 10 de marzo, está representado el día 28 de febrero.



**Figura 24.** RMSE medio de las parcelas de CT para el cultivo de la cebada en regadío en función del momento en el que se simulan los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.



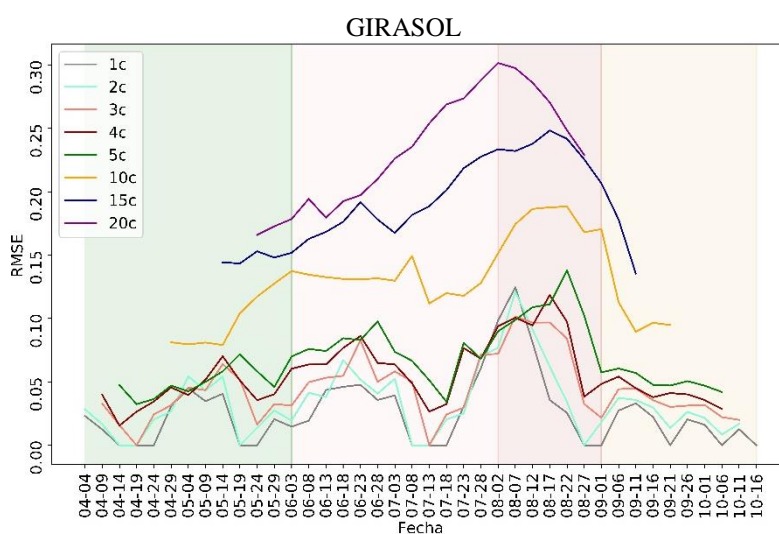
**Figura 25.** RMSE medio de las parcelas de CT para el cultivo del trigo blando en secano (centro) en función del momento en el que se simulan los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

Puede observarse que la curva del RMSE sigue una forma de campana cuando la duración de los huecos es más larga (15c y 20c), alcanzando los picos su máximo en el centro del periodo



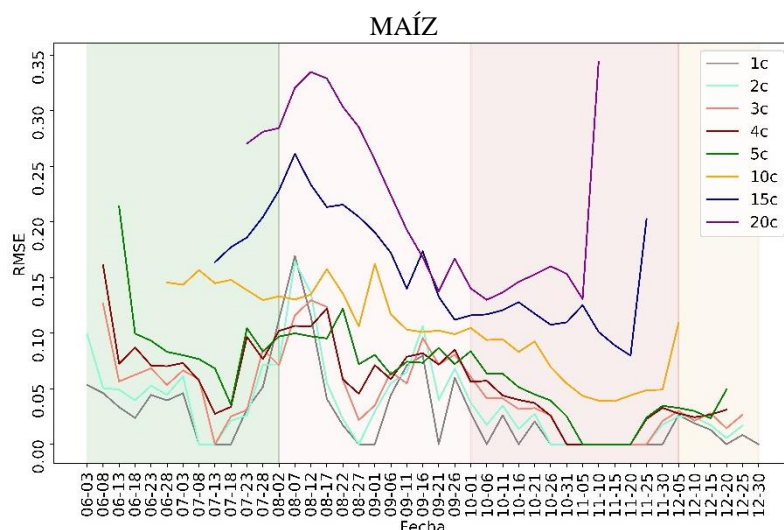
vegetativo del cultivo, coincidiendo con el final de la etapa de crecimiento del tallo y al principio del crecimiento de la inflorescencia en la cebada y el trigo (**Figura 24** y **Figura 25**). Como se ha visto en apartado 5.7, en estas fases es cuando se observa un mayor incremento en la respuesta de los índices de S2 a las variaciones del cultivo, por lo que la falta de datos en esas fechas se refleja en un mayor RMSE. Se observa también cómo a medida que aumenta la duración del hueco, el RMSE máximo se adelanta. Es decir, cuanto mayor es la duración del hueco, el RMSE afecta en fases más tempranas, mientras que, cuando la duración del hueco es menor, el RMSE se alcanza en fases más tardías. Una vez alcanzado el máximo, el RMSE vuelve a descender hasta llegar a valores mínimos en el período de llenado y madurez o cosecha. Esta tendencia se observa de manera más acentuada cuanto mayor es la duración de los huecos.

El girasol (**Figura 26**) muestra resultados muy similares a los cultivos de invierno, mostrando ese retraso en el pico máximo que también se ha observado en la **Figura 20**. El pico de RMSE máximo se adelanta a medida que la duración de los huecos es mayor, alcanzando el valor máximo hacia el final de la fase reproductiva, y no al final de la fase de llenado, como podría esperarse. En comparación con los RMSE obtenidos en los cultivos de invierno, tanto el girasol como el maíz (**Figura 27**) presentan valores superiores, indicando esto que, en este particular caso, el impacto de los huecos en los cultivos de verano es mayor que en los de invierno.



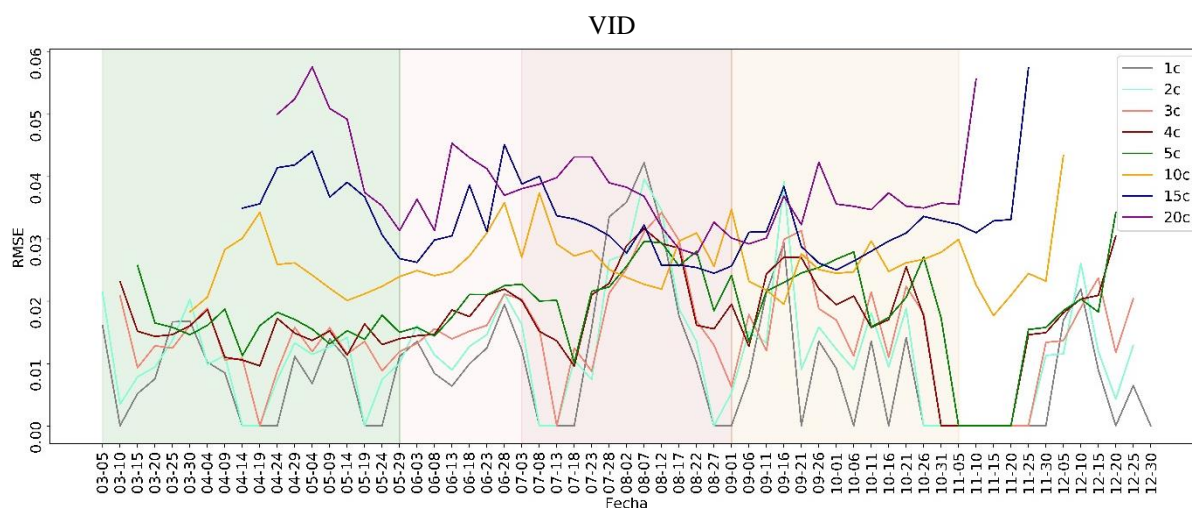
**Figura 26.** RMSE medio de las parcelas de CT para el cultivo del girasol en función del momento en el que se simulan los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

La evolución del RMSE del maíz a lo largo del período de estudio (**Figura 27**) sigue un comportamiento parecido a los cultivos recién comentados, pero con picos más acentuados a lo largo de todo el periodo. El periodo crítico vuelve a ser la fase reproductiva donde se sitúan los picos de RMSE máximos.



**Figura 27.** RMSE medio de las parcelas de CT para el cultivo del maíz en función del momento en el que se simulan los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

En cuanto a la vid (**Figura 28**), la evolución del RMSE a lo largo del periodo de estudio es bastante más regular que los cultivos de verano e invierno, alcanzando este un valor máximo hasta 5 veces menor (0,06), situándose los picos en el período vegetativo para duraciones de huecos largas (15 y 20c) y en el periodo reproductivo para huecos más cortos. Era de esperar que los errores de la vid fueran menores al resto de cultivo debido a que los valores de NDVI alcanzados en la serie temporal son mucho menores que otros cultivos.

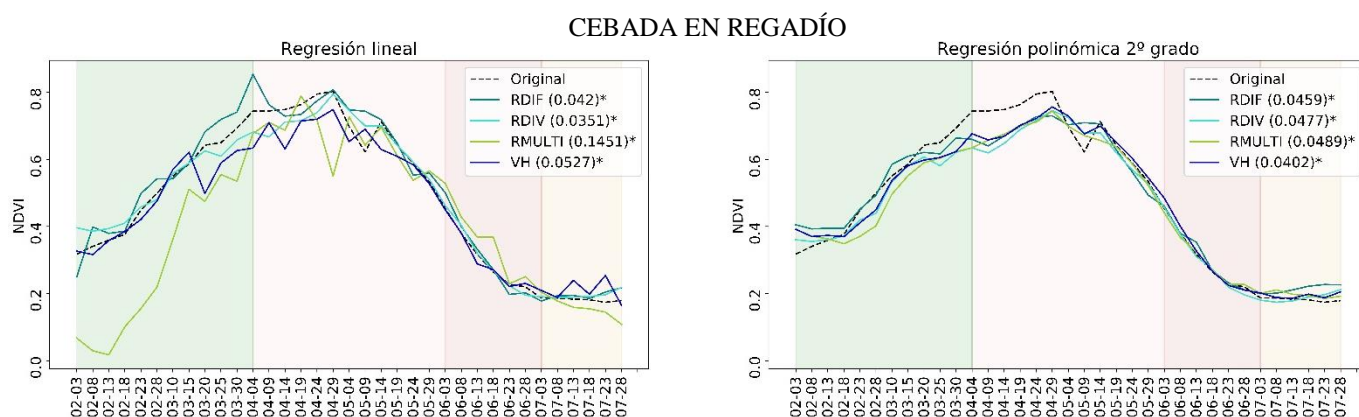


**Figura 28.** RMSE medio de las parcelas de CT para el cultivo de la vid, para una ventana temporal de 01/03/2019 – 31/12/2019, en función del momento en el que se simulan los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

## 5.9. Evaluación de las técnicas de relleno de huecos

### 5.9.1. Relleno a partir de datos de Sentinel-1

La **Figura 29** muestra las series de NDVI estimadas a partir de datos de índices de S1 mediante regresión lineal (izquierda) y regresión polinómica de 2º grado (derecha) para el cultivo de cebada, presentando en paréntesis el valor de RMSE entre los huecos reconstruidos (obtenidos a partir de las curvas estimadas) y los originales. A partir de las ecuaciones de regresión ajustadas se estimaron los valores de NDVI para las fechas con hueco. En el caso de la cebada en regadío, tanto en la regresión lineal como en la polinómica, la serie que presenta mayor RMSE frente a la original es aquella obtenida a partir del índice RMULTI, pudiendo esperar que la reconstrucción de la serie a partir de estos datos muestre peores resultados que los demás. Las series de referencia logradas a partir de los índices RDIF y RDIV presentan RMSE menores al ser obtenidas a partir de regresiones lineales, mientras que las calculadas a partir de los índices radar RMULTI y VH presentan errores bastante menores con la regresión polinómica (ver ANEXO V. SERIES DE NDVI DE REFERENCIA PARA EL RELLENO DE HUECOS A PARTIR DE DATOS DE SENTINEL-1 para consultar las series de referencia del resto de cultivos).



**Figura 29.** Serie temporal NDVI estimada obtenida a partir de datos de índices Sentinel-1 mediante regresión lineal (izquierda) y regresión polinómica de segundo grado (derecha) para el cultivo de la cebada en regadío. \*RMSE Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

En la **Tabla 19** se resumen los valores de RMSE analizados en la **Figura 29** para los cultivos de interés. Exceptuando para los cultivos de cebada y trigo en regadío, las series de referencia obtenidas a partir de una regresión polinómica de 2º grado presentan menor RMSE respecto a la serie original que las series producidas mediante regresión lineal. Para la cebada en secano, la serie de referencia con menor RMSE es aquella obtenida a partir del índice RDIV mediante regresión lineal (0,035), mientras que el mínimo RMSE para la cebada en regadío se obtiene a partir del índice VH mediante regresión polinómica (0,029). En el caso del trigo, el menor RMSE se obtiene con el índice VH, pero a partir de regresión lineal en regadío (0,025) y regresión polinómica en secano (0,025). Para el caso del girasol y el maíz, el mínimo RMSE se obtiene mediante la regresión polinómica, a partir de RDIF (0,019) y RDIV (0,021) respectivamente. Es

la vid el cultivo que menor RMSE presenta, obtenido a partir del índice RMULTI mediante regresión polinómica (0,017). En cuanto a los RMSE máximos, el índice que peor resultados muestra es el VH en regresión lineal en los cultivos de cebada en secano, girasol y vid, el RMULTI en la cebada en regadío y el maíz (regresión lineal), el RDIV en trigo en secano y el RDIF en regadío.

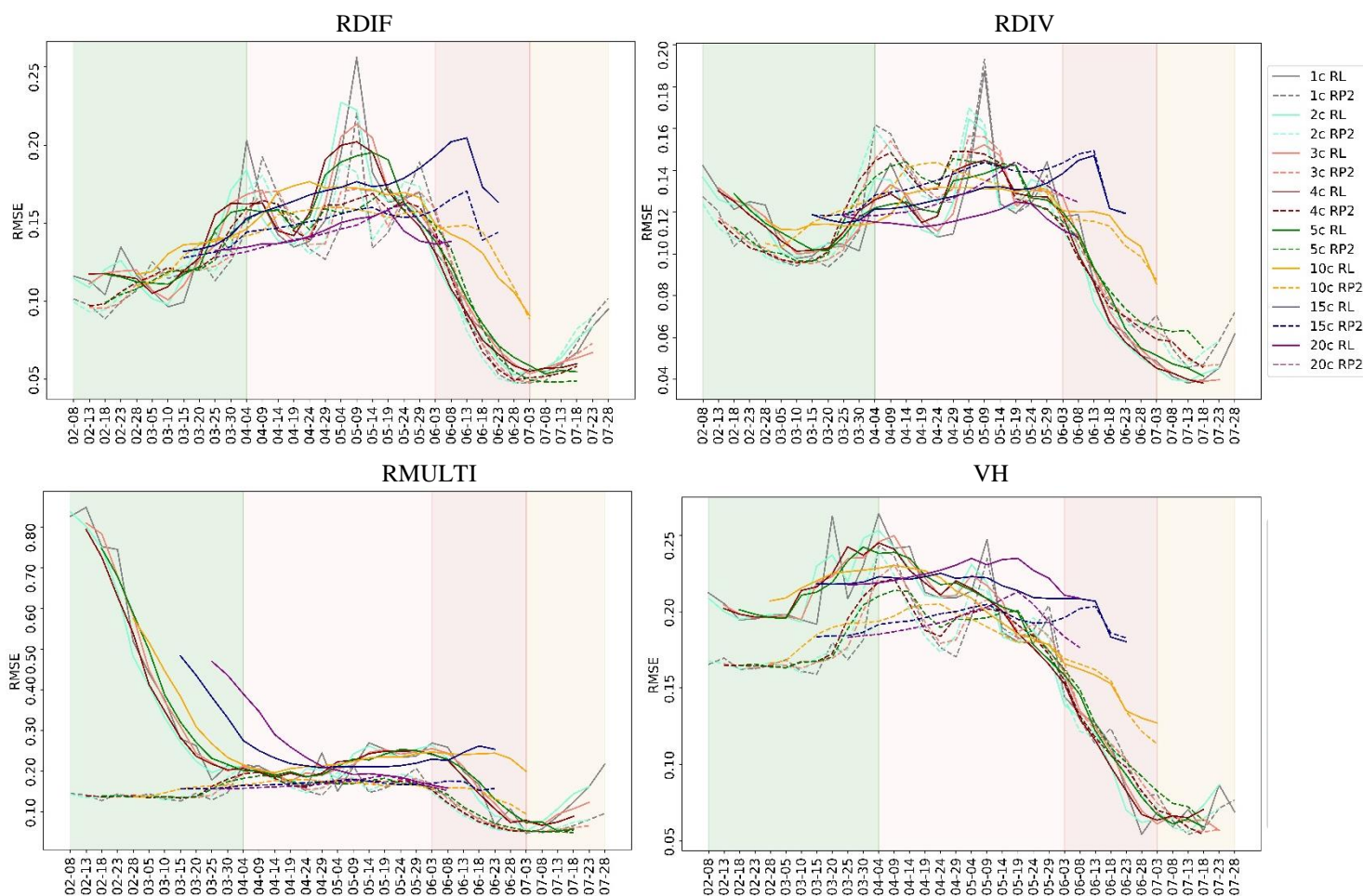
**Tabla 19.** Resumen de valores de RMSE entre la serie temporal original de NDVI y la serie de referencia de NDVI obtenida a partir de los diferentes índices radar mediante regresión lineal (RL) y regresión polinómica de 2º grado (RP2) para los cultivos de interés.

	RDIF		RDIV		RMULTI		VH	
	RL	RP2	RL	RP2	RL	RP2	RL	RP2
<b>CEBADA R.</b>	0,042	0,046	0,035	0,048	0,145	0,049	0,053	0,040
<b>CEBADA S.</b>	0,037	0,047	0,050	0,059	0,045	0,053	0,061	0,029
<b>TRIGO R.</b>	0,070	0,051	0,050	0,028	0,066	0,031	0,025	0,036
<b>TRIGO S.</b>	0,070	0,040	0,131	0,038	0,099	0,033	0,054	0,025
<b>GIRASOL</b>	0,036	0,019	0,045	0,035	0,093	0,036	0,093	0,068
<b>MAÍZ</b>	0,049	0,045	0,023	0,021	0,059	0,044	0,039	0,027
<b>VID</b>	0,035	0,017	0,044	0,020	0,026	0,017	0,052	0,026

La **Figura 30** muestra el RMSE medio entre los datos de NDVI de la serie original (STi) y los datos reconstruidos a partir de las series de referencia (STAR<sub>S1</sub>, obtenidas a partir de datos de índices de S1 de las parcelas del CE) de cebada en secano. Los índices RDIF, RDIV y VH muestran patrones de curvas similares, alcanzando los picos de mayor error en el período reproductivo, siendo el segundo el que mejores valores presenta (resultados acordes a los datos obtenidos en la **Tabla 19**). A diferencia de lo que podría esperarse, no se aprecia una clara tendencia de aumento en el error a medida que aumenta el tamaño del hueco, sino todo lo contrario. En el caso del índice RDIF, se obtienen picos de error superiores para una duración de huecos de 1c, disminuyendo estos picos máximos a medida que aumenta la duración del hueco, hasta las 5c. Los huecos de 10c, 15c y 20c siguen una curva más suavizada, con picos menos pronunciados. En cuanto a las diferencias entre los resultados obtenidos mediante regresión lineal y polinómica, de forma general, la mayoría de duraciones de huecos presentan errores menores en la regresión polinómica que en la lineal, tal y como se muestra en la **Tabla 19**. El índice RDIV sigue patrones similares a los recién comentados, a diferencia de que, tal y como se podía prever por los datos obtenidos en **Tabla 19**, se obtienen errores menores a partir de regresión lineal. Las curvas de RMSE presentadas a partir del índice VH, aunque sean también similares a los dos índices comentados, tienen un ligero adelanto, presentando los picos al principio del período reproductivo. Los errores del índice RMULTI siguen un patrón bastante diferente: mediante regresión lineal se alcanzan los picos máximos en el período vegetativo, alcanzando valores máximos de hasta 0,80, viéndose estos reducidos a lo largo del ciclo fenológico del cultivo. Los errores obtenidos a partir de regresión polinómica siguen una tendencia más plana, alcanzando los

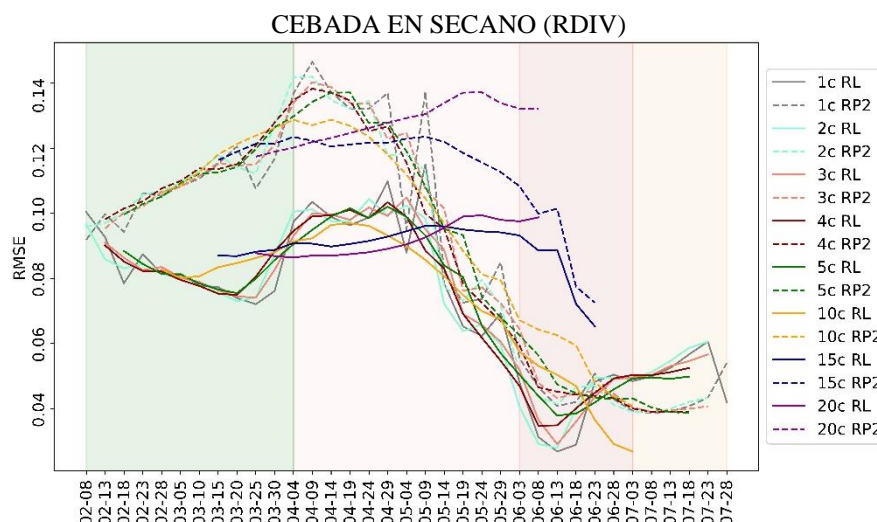


máximos en el período reproductivo, siendo estos menores que los alcanzados en la regresión lineal en esta misma fase.



**Figura 30.** RMSE medio entre la serie de NDVI original ( $ST_i$ ) y las series reconstruidas ( $STAR_{Si}$ ) a partir de índices radar mediante regresiones lineales (RL, línea continua) y polinómicas (RP2, línea discontinua) para la cebada en regadío, en función de la duración de los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

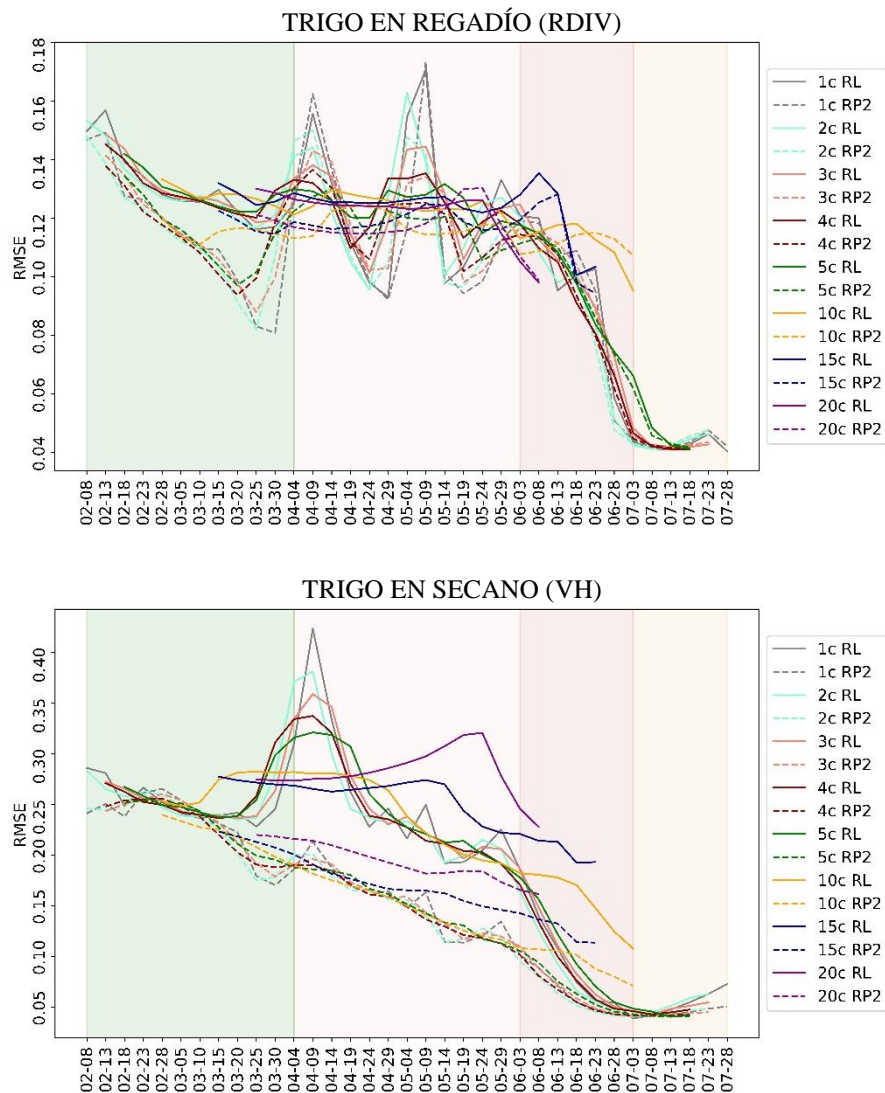
En el caso de la cebada en secano (**Figura 31** con los resultados del índice que mejores RMSE presenta), de acuerdo con los resultados analizados en la **Tabla 19**, podría esperarse obtener los mejores resultados de RMSE a partir del índice VH mediante regresión polinómica. Sin embargo, los errores menores se obtienen a partir del índice RDIV mediante regresión, tal y como la cebada en regadío, aunque alcanzando picos menores (máximos de 0,12 en secano vs. máximos de 0,18 en regadío). Los comportamientos de las curvas obtenidas son bastante similares a las descritas para la cebada en regadío, volviendo a ser el período más crítico la fase reproductiva, aunque en este caso no se observe una diferencia tan clara de disminución del error a medida que la duración del hueco aumenta.



**Figura 31.** RMSE medio entre la serie de NDVI original (STi) y las series reconstruidas (STARSi) a partir del índice radar RDIV mediante regresiones lineales (RL) y polinómicas (RP2) para la cebada en secano, en función de la duración de los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

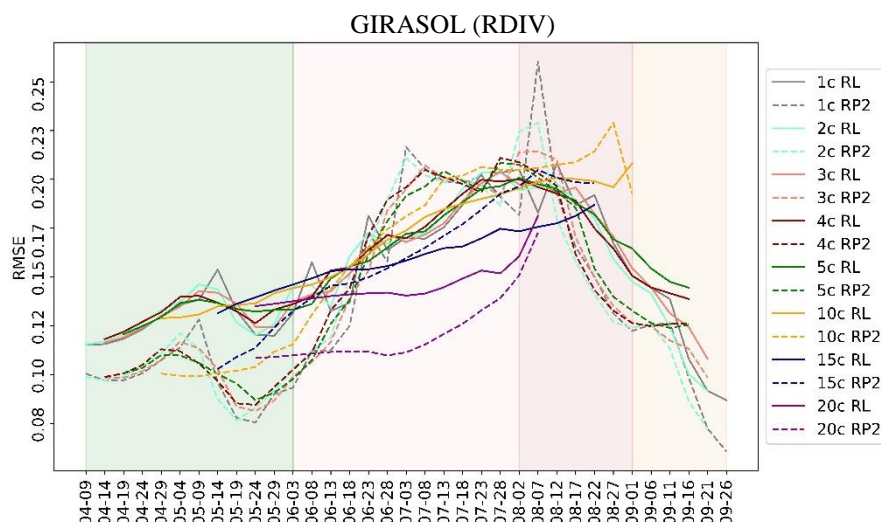
El trigo en regadío (**Figura 32** arriba) presenta mejores resultados de relleno al utilizar regresiones polinómicas en los cuatro índices radar, observándose mayor diferencia entre las dos opciones en el periodo vegetativo. De nuevo es el índice RDIV el que presenta valores de RMSE menores (picos máximos que no alcanzan errores de 0,18) y, por lo tanto, curvas rellenas finales mejores. Vuelve a observarse el mismo patrón, donde el período crítico (los valores máximos de RMSE) se encuentra en la fase reproductiva, en los huecos de duraciones de 1 a 3 capturas. Se observan de nuevo tendencias más planas en las duraciones de huecos superiores (de 10 a 20c). En el trigo en secano (**Figura 32** abajo) se obtienen resultados diferentes. Aunque los errores obtenidos a partir de regresión lineal presentan una curva en forma de campana como se ha ido observando para otros casos, las curvas de RMSE obtenidas a partir de regresiones polinómicas siguen una tendencia descendente, siendo de nuevo RDIV el índice a partir del cual se extraigan mejores curvas.





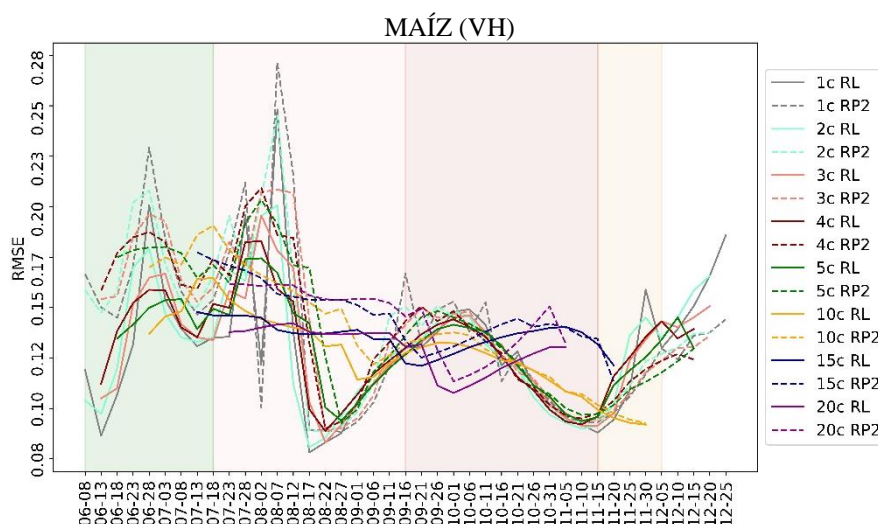
**Figura 32.** RMSE medio entre la serie de NDVI original (STi) y las series reconstruidas (STAR<sub>Si</sub>) a partir del índice radar RDIV mediante regresiones lineales (RL) y polinómicas (RP2) para el trigo en regadío (arriba) y en secano (abajo), en función de la duración de los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

Para el caso del girasol (**Figura 33**), la curva de RMSE medio ha seguido la tendencia esperada, viéndose atrasados los máximos hasta el periodo de llenado y madurez, tal y como se observaba en el apartado 5.7. Se puede observar como los errores obtenidos mediante regresión lineal son menores que los obtenidos a partir de regresión polinómica para la mayoría de huecos. En cuanto al índice radar a partir del cual se obtienen los mejores resultados, vuelve a ser el RDIV el que obtiene errores menores. Vuelve a ser el hueco de 1c el que presente los picos máximos, seguidos de huecos de 2 y 3c en el periodo de llenado. Exceptuando por el pico máximo, la tendencia seguida por los huecos de 1c a 5c es bastante similar, tanto en los RMSE obtenidos a partir de regresiones lineales como polinómicas. Son los huecos de 10c a 20c los que vuelven a presentar una tendencia más plana, siendo los últimos los que menor error presentan entre los tres.



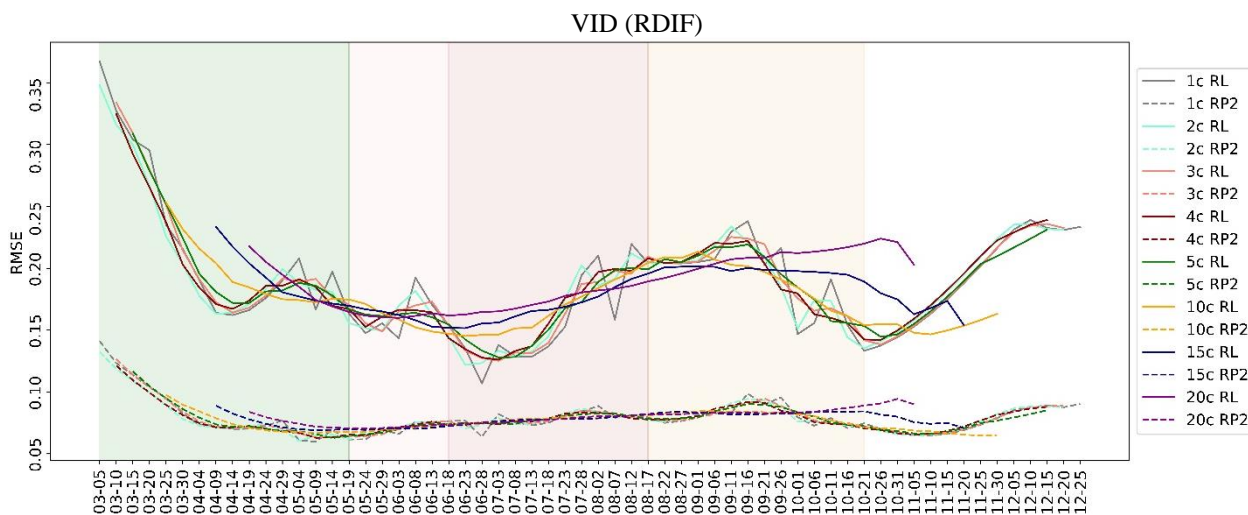
**Figura 33.** RMSE medio entre la serie de NDVI original ( $ST_i$ ) y las series reconstruidas ( $STAR_{S1}$ ) a partir los índices radar RDIF y RDIV mediante regresiones lineales (RL) y polinómicas (RP2) para el girasol, en función de la duración de los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

El maíz presenta unas curvas de RMSE ligeramente diferentes a las que se han ido viendo. Se obtiene un primer pico en el periodo vegetativo seguido de un pico máximo en el período reproductivo, a partir del cual comienza un descenso, alcanzando el mínimo también en el periodo reproductivo, para los huecos de 1c a 5c. Los huecos de 10c a 20c vuelven a seguir una tendencia más suavizada, pero, esta vez, descendente. En este caso, la diferencia entre la regresión lineal y la polinómica no es tan marcada, presentando la segunda errores superiores. Son las curvas obtenidas a partir del índice VH (**Figura 34**) las que mejores resultados presentan.



**Figura 34.** RMSE medio entre la serie de NDVI original ( $ST_i$ ) y las series reconstruidas ( $STAR_{S1}$ ) a partir el índice radar VH mediante regresiones lineales (RL) y polinómicas (RP2) para el maíz, en función de la duración de los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

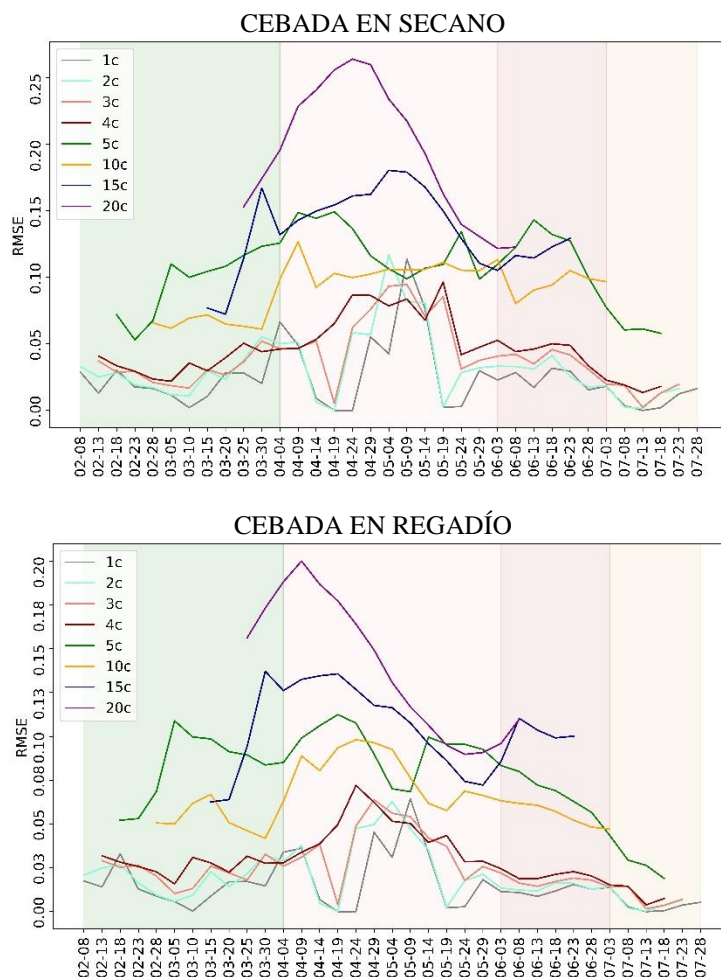
En el caso del cultivo de la vid (**Figura 35**), se observan curvas diferencias bastante notables entre los resultados obtenidos a partir de regresión lineal y polinómica, siendo esta última la que presenta errores bastante inferiores. En este caso, los menores errores se obtienen a partir del índice RDIF, con máximas que no llegan a 0,15 en el período vegetativo (donde se alcanza el pico máximo). En las siguientes fases, el error se mantiene bastante constante entre valores de 0,05 y 0,10.



**Figura 35.** RMSE medio entre la serie de NDVI original ( $ST_i$ ) y las series reconstruidas ( $STAR_{ST_i}$ ) a partir del índice radar RDIF mediante regresiones lineales (RL) y polinómicas (RP2) para la vid, en función de la duración de los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

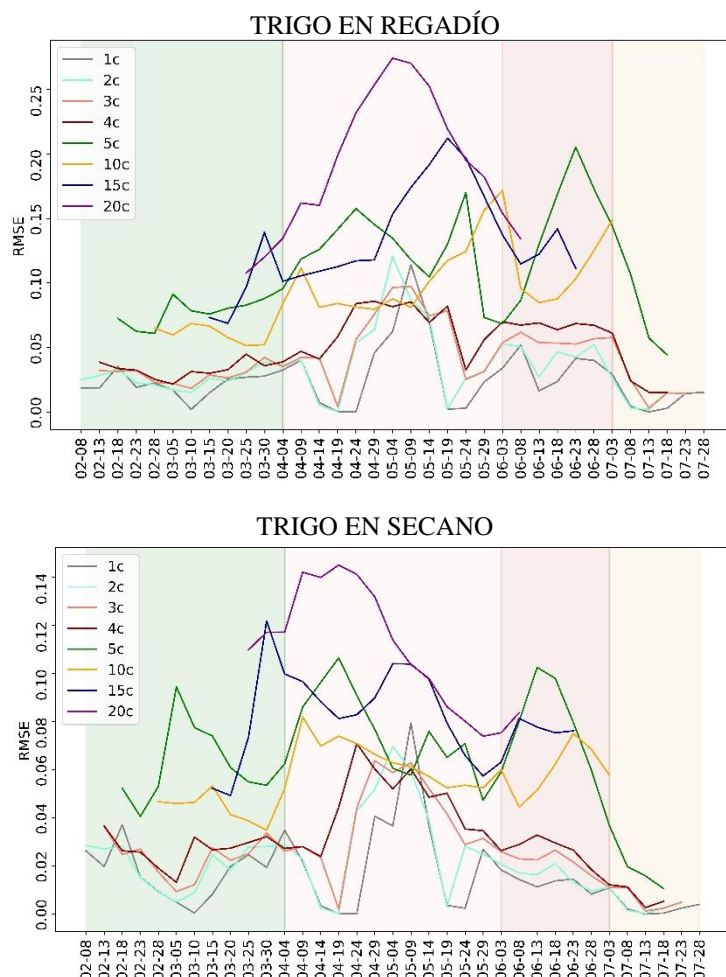
### 5.9.2. Relleno a partir de datos de la propia serie de Sentinel-2

La **Figura 36** muestra los valores de RMSE medios a lo largo del ciclo temporal del cultivo obtenidos a partir de la diferencia de los datos de NDVI de la serie temporal original (STi5) y los datos de NDVI las series reconstruidas a partir de las propias series de Sentinel-2 (STAR<sub>S2</sub>) para la cebada. La curva sigue un patrón esperado, alcanzando de nuevo los picos máximos de RMSE en la fase reproductiva (coincidiendo en fase con los errores obtenidos a través de la técnica de relleno a partir de datos de Sentinel-1), con valores superiores a 0,25 para una duración de hueco de 20 capturas para la cebada en regadío y de 0,20 en secano, volviendo a descender. Vuelve a observarse el adelantamiento de los picos en el secano, tal y como se ha analizado en el apartado 5.7, mostrando este último errores menores que la cebada en regadío para las duraciones de huecos analizadas. En este caso se observa una clara tendencia en el aumento del error a medida que aumenta la duración del hueco, a diferencia de lo observado en las curvas RMSE obtenidas a partir de datos de Sentinel-1 (ver apartado 5.9.1).



**Figura 36.** RMSE medio entre la serie de NDVI original (STi5) y las series reconstruidas a partir de la propia serie de Sentinel-2 (STAR<sub>S2</sub>) para la cebada en regadío (arriba) y secano (abajo), en función de la duración de los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

En cuanto al trigo (**Figura 37**), se observa que presenta patrones similares a la cebada, con RMSE menores en secano (0,14) pero mayores en regadío (0,25) en los huecos de duración de 20 capturas, obtenidas en el periodo reproductivo. A diferencia de lo observado en las curvas de RMSE obtenidas a partir de datos de Sentinel-1 (ver apartado 5.9.1), donde las curvas de huecos cortos y largos no distaba de manera significativa, en este caso del trigo (y los demás cultivos) se observa que hay mayor diferencia en los valores de RMSE de huecos cortos y huecos largos.

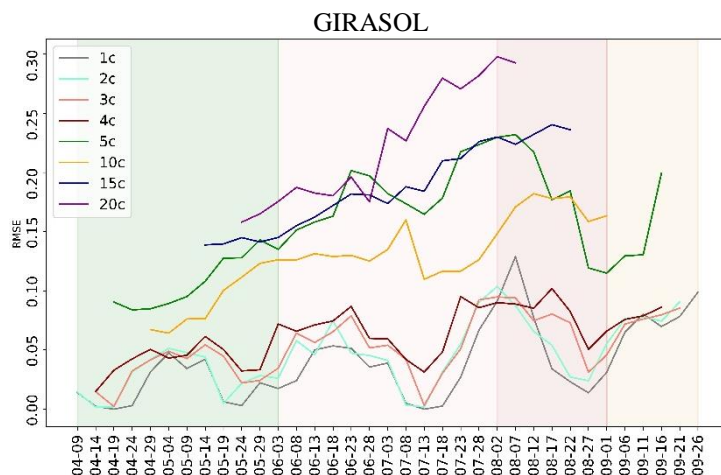


**Figura 37.** RMSE medio entre la serie de NDVI original (STi5) y las series reconstruidas a partir de la propia serie de Sentinel-2 (STAR<sub>s2</sub>) para el trigo en regadío (arriba) y secano (abajo), en función de la duración de los huecos.

Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

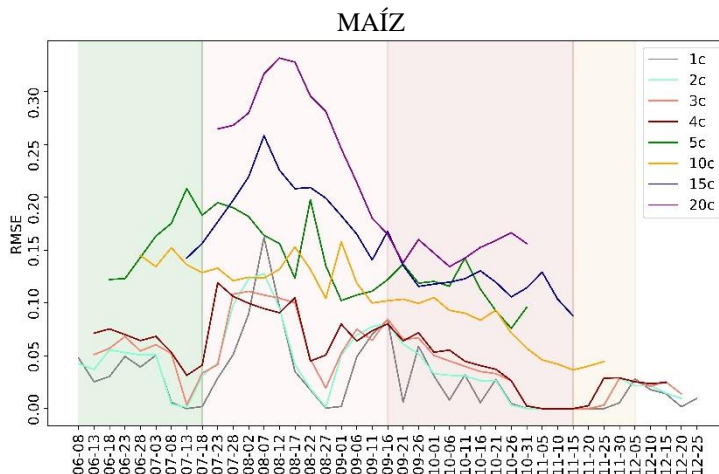
Como se ha ido observando a lo largo del trabajo, el girasol (**Figura 38**) vuelve a presentar los máximos de RMSE en el periodo de llenado, alcanzando valores de hasta 0,30 para una duración de huecos de 20 capturas. Puede volver a verse cómo las curvas de RMSE de los huecos de duración entre 1 y 4c varían entre RMSE de 0,00 y 0,10 mientras que, los huecos de duraciones más largas, presentan errores bastante más altos, superando en la mayoría de los casos errores de 0,10.





**Figura 38.** RMSE medio entre la serie de NDVI original (STi5) y las series reconstruidas a partir de la propia serie de Sentinel-2 (STAR<sub>S2</sub>) para el girasol, en función de la duración de los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

El maíz (**Figura 39**) vuelve a presentar valores de RMSE superiores en el periodo reproductivo, superando un error de 0,28 en huecos de 20 capturas, reduciéndose a medida que avanza el ciclo del cultivo. Vuelven a obtenerse errores menores para duraciones de huecos de 1 a 4 capturas, siendo el hueco de 1c el que presenta un pico máximo más pronunciado, estando comprendido estos valores de nuevo entre 0 y 0,10.

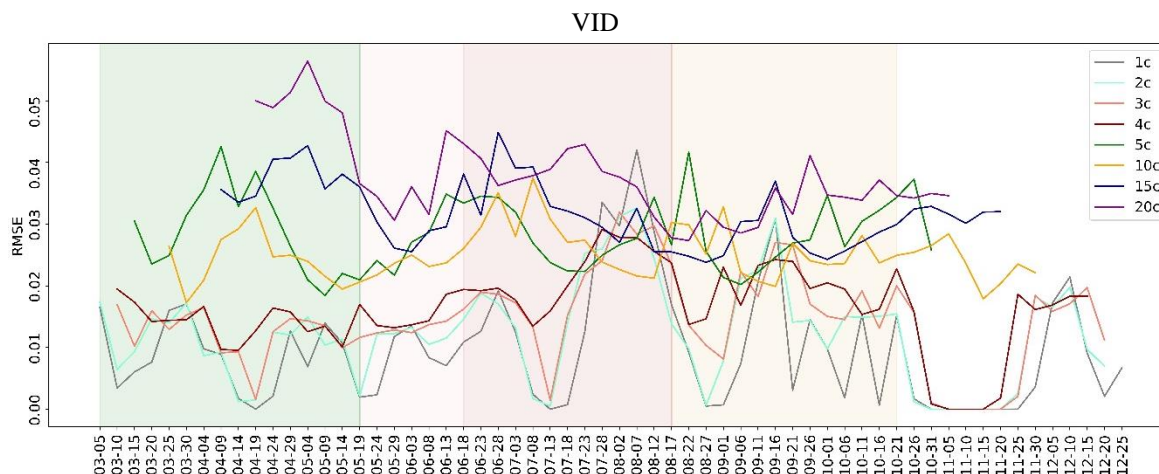


**Figura 39.** RMSE medio entre la serie de NDVI original (STi5) y las series reconstruidas a partir de la propia serie de Sentinel-2 (STAR<sub>S2</sub>) para el maíz, en función de la duración de los huecos: 1c- 5c (izquierda) y 5c – 20c (derecha), en función de la duración de los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

Finalmente, la vid (**Figura 40**) vuelve a presentar curvas más regulares, sin picos tan pronunciados como se ha ido viendo en los demás cultivos. En este caso, las diferencias en RMSE entre los huecos de diferente no varía tanto, sobrepasando ligeramente el pico máximo de toda la serie un error de 0,05. Tal y como se ha comentado en el apartado 5.8, los valores de NDVI que



alcanza la vid son bastante menores que los otros cultivos. Es por ello que de nuevo se vuelven a obtener valores de RMSE bastante menores.

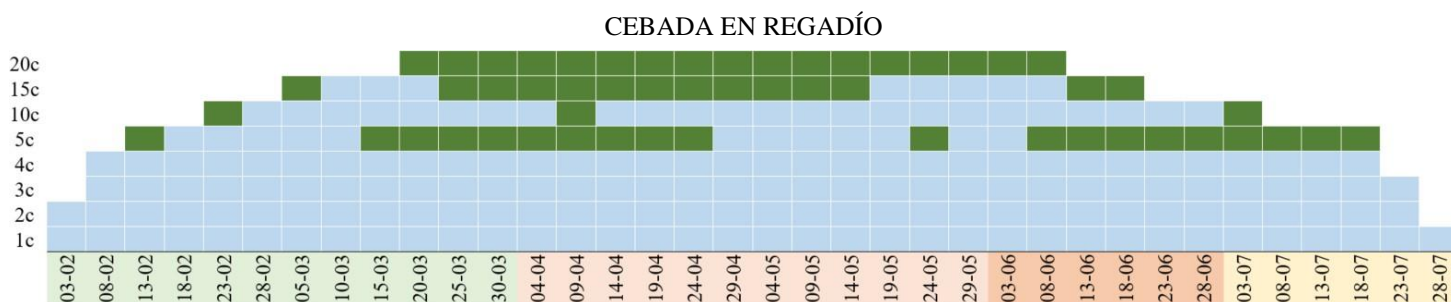


**Figura 40.** RMSE medio entre la serie de NDVI original (STi5) y las series reconstruidas a partir de la propia serie de Sentinel-2 (STARs2) para el maíz, en función de la duración de los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

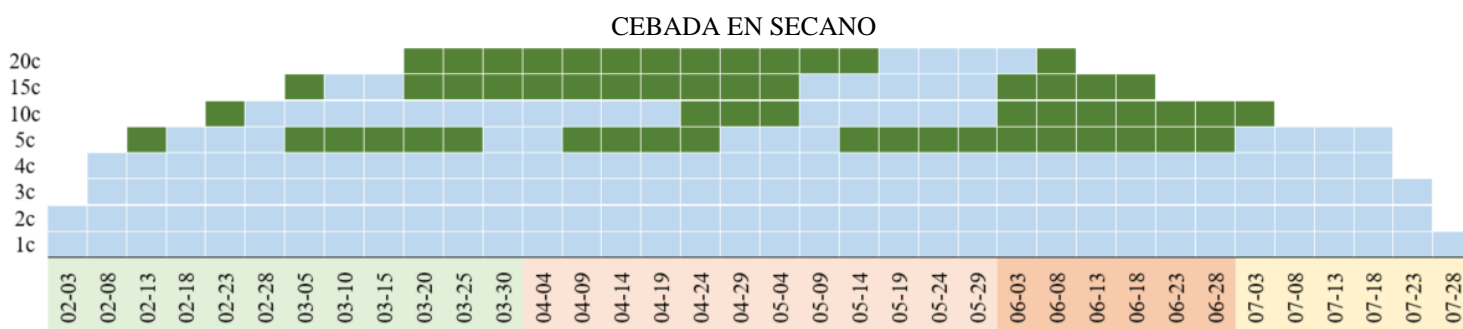
### 5.10. Comparación entre técnicas de relleno

De forma general para los siete cultivos se ha observado que, al aplicar la técnica de relleno de huecos a partir de datos de Sentinel-2, cuando la duración de los huecos era corta (de 1 a 4 capturas), el error obtenido era bastante menor que el RMSE calculado en huecos de duración mayor (de 5 a 20 capturas). Sin embargo, esta diferencia no se ha observado de manera tan clara en la técnica de relleno a partir de datos de Sentinel-1.

La **Figura 41** muestra un resumen de la técnica que mejores resultados presenta (menor RMSE) en función de la fecha en la que ocurre el hueco y la duración de este, correspondiendo el color verde a la técnica de relleno a partir de datos Sentinel-1 (en este caso mediante regresión lineal, a partir del índice RDIV) y el color azul a la técnica de relleno a partir de datos Sentinel-2. Puede observarse que, para duraciones de huecos de 1 a 5 capturas, la técnica de relleno de datos a partir de Sentinel-2 es la que presenta mejores resultados, mientras que, para duraciones de 20c, es a partir de Sentinel-1 cuando se obtienen mejores resultados, para todas las fechas. Para el resto de huecos (5c, 10c y 15c), la técnica varía en función de la fecha en la que ocurre el hueco. La cebada en secano (**Figura 42**) presenta los mismos resultados que en regadío en huecos de 1 a 4 capturas, presentando cambios en huecos de mayor duración.

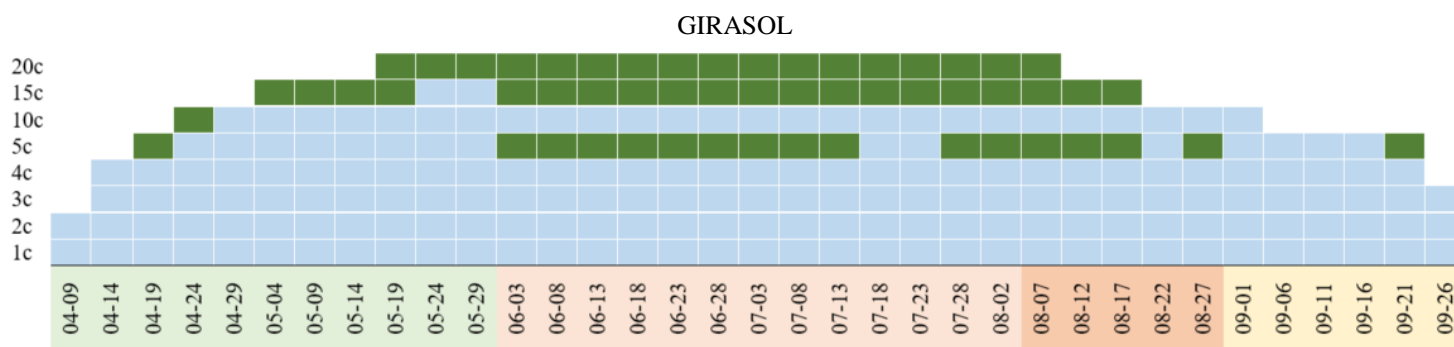


**Figura 41.** Resumen de las técnicas que presentan mejores resultados (menor RMSE) en función de la fecha en la que ocurre el hueco y la duración de este, correspondiendo el color azul a Sentinel-2 y el color verde a Sentinel-1 (regresión lineal, RDIV), para el cultivo de cebada en secano. Fechas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.



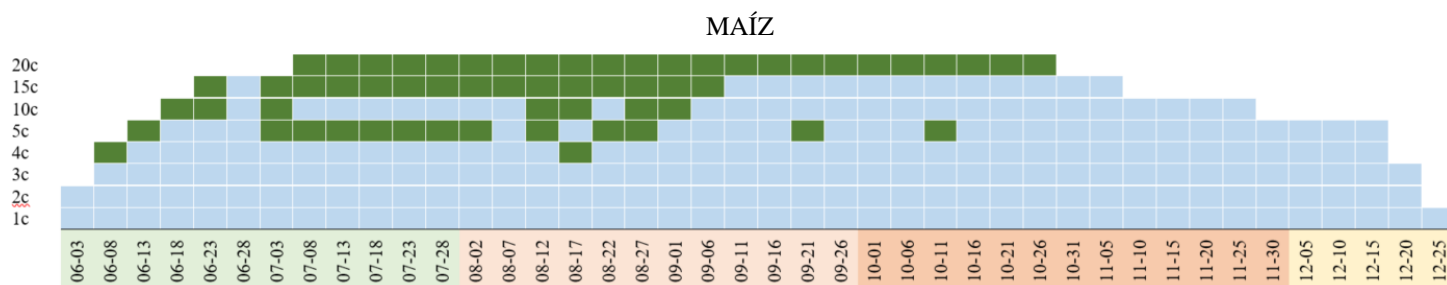
**Figura 42.** Resumen de las técnicas que presentan mejores resultados (menor RMSE) en función de la fecha en la que ocurre el hueco y la duración de este, correspondiendo el color azul a Sentinel-2 y el color verde a Sentinel-1 (regresión lineal, RDIV), para el cultivo de cebada en secano. Fechas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

En el girasol (**Figura 43**) vuelve a obtenerse mejores resultados mediante la técnica de relleno a partir de datos de Sentinel-2 en huecos de 1 a 4 capturas, mientras que la técnica óptima variará en función de la duración del hueco y la fecha en el que este ocurre. En este caso se obtiene que, para huecos de 15 (a excepción de dos fechas) y 20c, es la técnica de relleno a partir de datos de Sentinel-1 la que presenta menores errores.



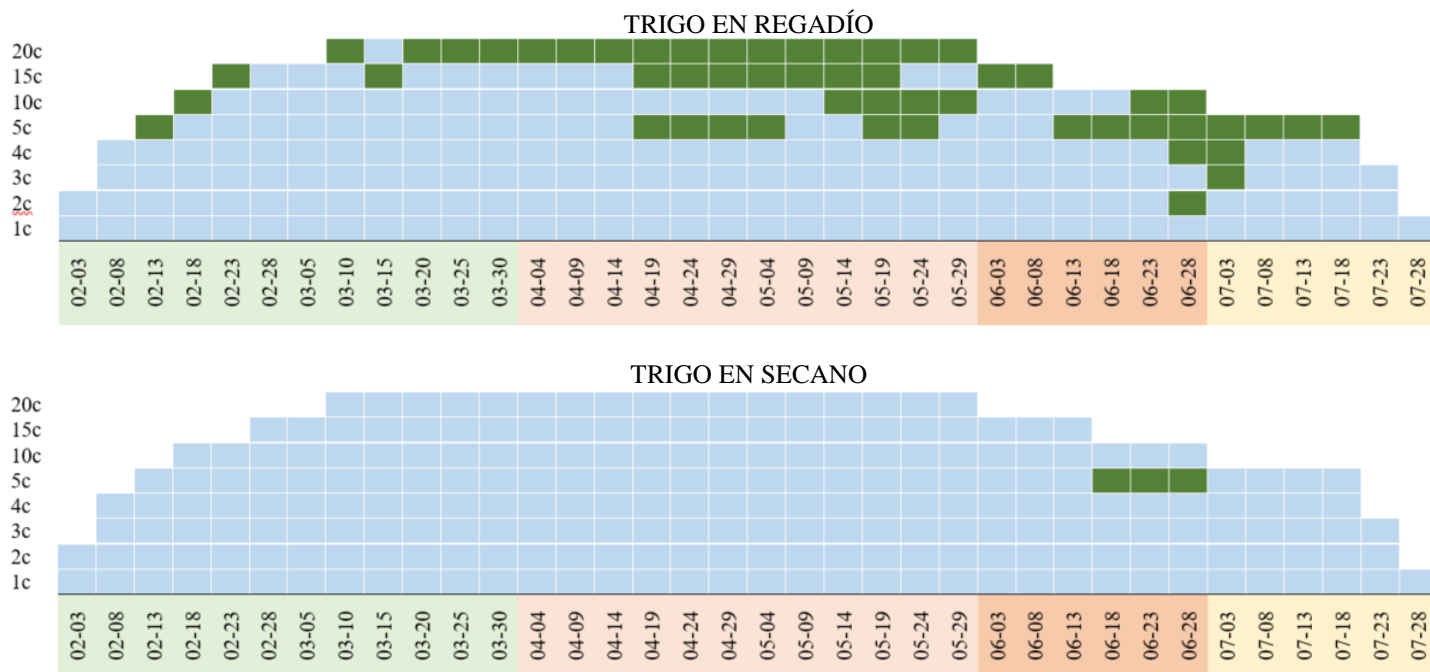
**Figura 43.** Resumen de las técnicas que presentan mejores resultados (menor RMSE) en función de la fecha en la que ocurre el hueco y la duración de este, correspondiendo el color azul a Sentinel-2 y el color verde a Sentinel-1 (regresión lineal, RDIV), para el cultivo de girasol. Fechas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

En el cultivo del maíz (**Figura 44**) se vuelve a observar patrones similares, con la técnica de relleno a partir de datos Sentinel-2 como mejor opción en huecos de 1 a 4 capturas. Se observa de nuevo cómo a medida que aumenta la duración de los huecos, en función de la fecha donde ocurran, la técnica de relleno a partir de datos Sentinel-1 resulta ser mejor opción en algunos casos.



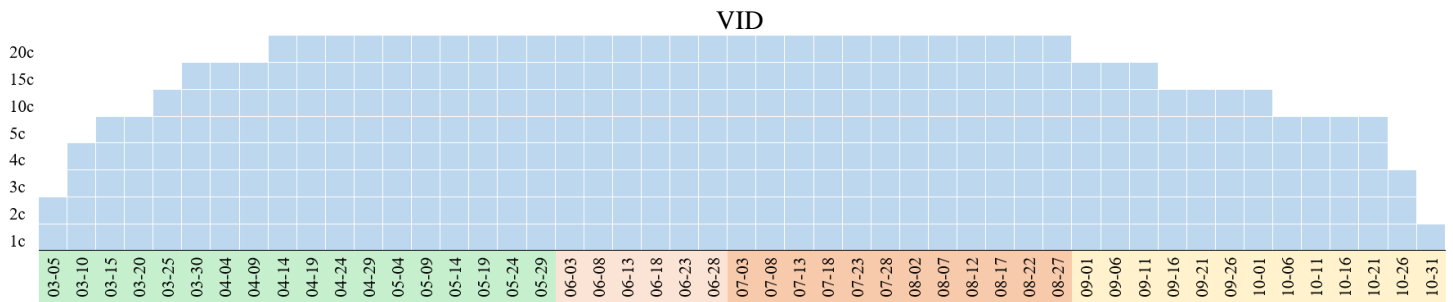
**Figura 44.** Resumen de las técnicas que presentan mejores resultados (menor RMSE) en función de la fecha en la que ocurre el hueco y la duración de este, correspondiendo el color azul a Sentinel-2 y el color verde a Sentinel-1 (regresión lineal, VH), para el cultivo de maíz. Fechas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

Mientras que el trigo en regadío ha presentado resultados muy similares a los que se han ido observando para el resto de cultivos (**Figura 45**, arriba), el trigo en secano presenta resultados bastante diferentes, presentando la técnica de relleno a partir de datos Sentinel-2 como mejor opción para casi todos los casos.



**Figura 45.** Resumen de las técnicas que presentan mejores resultados (menor RMSE) en función de la fecha en la que ocurre el hueco y la duración de este, correspondiendo el color azul a Sentinel-2 y el color verde a Sentinel-1 (regresión lineal, RDIV en regadío, regresión lineal VH en secano), para el cultivo de trigo. Fechas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

Finalmente, tal y como era de esperar debido a los errores analizados en los apartados 5.9.1 y 5.9.2, el relleno de huecos en el caso de la vid será mejor realizarlo a partir de datos de la propia serie de Sentinel-2, tal y como muestra la **Figura 46**.



**Figura 46.** Resumen de las técnicas que presentan mejores resultados (menor RMSE) en función de la fecha en la que ocurre el hueco y la duración de este, correspondiendo el color azul a Sentinel-2 y el color verde a Sentinel-1 (regresión polinómica, RDIF), para el cultivo de la vid. Fechas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

## 6. CONCLUSIONES

El presente trabajo tenía como objetivo principal analizar técnicas de relleno de huecos en series temporales de índices de vegetación obtenidos por Sentinel-2. Para ello, se han evaluado técnicas de relleno con datos procedentes de Sentinel-1 y con datos de la propia serie temporal de Sentinel-2 en diferentes cultivos de interés, analizando su implementación en función del momento y la duración de los huecos generados de forma artificial en las series temporales de NDVI. Además, también se han analizado las diferencias en el cálculo de índices de vegetación a nivel de pixel y a nivel de objeto (parcela) en las series temporales de S1 y S2, la correlación entre los índices de S1 y S2 calculados, y se ha interpretado el comportamiento reflexivo y retrodispersivo de los 7 cultivos de interés de la zona de estudio a partir de las series temporales de los índices calculados.

Por un lado, los resultados han mostrado que realizar el cálculo de los índices a nivel de pixel o a nivel de objeto (parcela) dependerá del índice y del cultivo (y, por ende, cubierta) que se analice. En este caso, realizar el cálculo de los índices ópticos NDVI, NDVI<sub>re</sub>, SAVI y SeLI en cualquiera de los dos niveles no debería suponer diferencias significativas en los resultados, siendo el índice MCARI el que peor resultados muestra. Es el cultivo del girasol el que menor error y mayor correlación presenta entre las dos opciones de cálculo. En cuanto a los índices radar, de forma general se muestran más sensibles que los ópticos, podría decirse que el índice VDDPI presentaría resultados similares calculándose este de las dos formas comentadas, mientras que el cálculo del índice IDPDD a nivel de parcela implicaría cometer grandes errores.

Por otro lado, para este caso de análisis de 7 cultivos, se ha observado que varios de los índices (tanto ópticos como radar) planteados presentan comportamientos (y, por lo tanto, información) similar, pudiendo estos agruparse, sin perder información relevante. En este caso, la selección inicial de 10 índices ópticos se redujo en 5 índices (MCARI, MTCI, NDMI, NDVI y NDVI<sub>re</sub>) y los 9 índices radar en 4 (RDIF, RDIV, RMULTI, VH).

En cuanto a las unidades de cálculo de los índices radar, los resultados han mostrado que las diferencias de correlaciones han sido poco significativas, siendo algo mayor cuando los índices radar se calculan en dB, a excepción del índice óptico NDVI<sub>re</sub>, que muestra los mismos coeficientes de correlación en ambos casos y que, además, son algo superiores que con otros índices de vegetación ópticos. Puede concluirse, en este caso, que la unidad en la que se realice el cálculo de los índices no debería implicar grandes diferencias en los resultados. Aplicar ventanas temporales ajustadas al ciclo fenológico de cada cultivo ha supuesto una mejora en la correlación entre índices ópticos y radar en la mayoría los cultivos analizados, con excepción de los cultivos de cebada en secano y trigo en secano. El descenso de la correlación tras la aplicación de la ventana temporal puede deberse, entre otras cosas, a la influencia que pueda tener un cultivo

precedente o posterior en las mismas parcelas. Por ello, se concluye que procede trabajar con ventanas temporales específicas con el fin de que, las correlaciones reflejan únicamente el desarrollo de cada cultivo. En cuanto al suavizado de curvas, ha demostrado mejorar de forma considerable las correlaciones entre las series temporales de índices ópticos y radar, siendo estas más altas cuando el tamaño de la ventana del filtro de Savitzky-Golay era mayor (31 días). El suavizado de curvas actualmente está siendo muy utilizado en los estudios de las trayectorias estacionales de la vegetación, en los proyectos de alta resolución en los que se analiza tanto la fenología como la productividad de esta (*Copernicus Land Monitoring Service*).

En lo que respecta a los comportamientos reflexivos y retrodispersivos de los cultivos, se ha visto que los cultivos de invierno siguen un patrón muy similar, mostrando cierto adelanto en esta tendencia el secano frente al regadío. Los 5 grupos de índices ópticos han mostrado curvas muy similares entre sí, representando de manera bastante clara las variaciones del cultivo a lo largo del ciclo fenológico en cada cultivo, reflejando una tendencia ascendente desde el comienzo de la germinación-periodo vegetativo hasta la mitad periodo reproductivo (en cultivos de invierno) / final del periodo de llenado (girasol) / principio del periodo de llenado (maíz) / periodo de llenado (vid), seguida de una tendencia descendente hasta el periodo de cosecha. El comportamiento retrodispersivo de los cultivos es un poco más complejo, reflejando el grupo VH la retrodispersión de volumen y el grupo RDIF el comportamiento retrodispersivo de VV.

En relación con los huecos en las series temporales de NDVI, los resultados han mostrado que existe una relación directa entre la duración de los huecos y el error que supone no disponer de esos datos en la serie original. Se ha observado también que existe relación entre el momento en el que ocurren dichos huecos y el error frente a la serie original, siendo mayor este en los momentos más críticos del ciclo fenológico de cada cultivo (periodo reproductivo). En el presente trabajo se ha decidido aplicar las técnicas sobre el índice NDVI en vez del NDVI<sub>re</sub>, aunque este último presentara mejores resultados, por un lado, debido a su amplio uso en la agronomía, y, por otro, debido a que no todos los satélites tienen la capacidad de capturar en las bandas necesarias para realizar el cálculo del índice NDVI<sub>re</sub>, pudiendo ser interesante dedicarle más atención a dicho índice en futuros trabajos de investigación.

Finalmente, en lo que respecta a las técnicas de relleno de huecos, se ha observado, por un lado, que los periodos en los que se producen los RMSE máximos para cada cultivo coinciden en ambas técnicas, con excepción del cultivo de la vid. Por otro lado, se ha obtenido que la idoneidad de cada una de las técnicas dependerá de cada cultivo, de la fecha en la que ocurra el hueco y de la duración de este. De norma general, para la mayoría de cultivos, para huecos de duración de 1 a 4 capturas, será la técnica de relleno de huecos a partir de datos de Sentinel-2 la que presente menores errores, mientras que, cuando el hueco sea más largo, dependerá de dónde ocurra este.



Para la cebada en secano será mejor utilizar la técnica a partir de datos de Sentinel-1 (regresión lineal e índice RDIV) en huecos de 20 capturas, independientemente de la fecha en la que ocurra el hueco, mientras que para huecos de 5 a 15 capturas dependerá de la fecha en la que el hueco ocurra. El girasol, el trigo en regadío y el maíz siguen un comportamiento muy similar, aunque para este último, en la técnica de relleno de huecos a partir de Sentinel-1, los mejores resultados se obtienen mediante el índice VH. El trigo en secano y la vid, al presentar errores muy bajos en la técnica de relleno de Sentinel-2, es esta técnica la que ofrece mejores resultados, independientemente de la duración del hueco y de la posición de este en la serie de estudio (comparados con la técnica de relleno de huecos a partir de Sentinel-1 mediante regresión lineal y VH en el trigo en secano y regresión polinómica y RDIF en la vid).

## 7. BIBLIOGRAFÍA

- Atzberger, C., & Eilers, P. H. C. (2011). A time series for monitoring vegetation activity and phenology at 10-daily time steps covering large parts of South America. *International Journal of Digital Earth*, 4(5), 365–386. <https://doi.org/10.1080/17538947.2010.505664>
- Beck, P. S. A., Atzberger, C., Høgda, K. A., Johansen, B., & Skidmore, A. K. (2006). Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI. *Remote Sensing of Environment*, 100(3), 321–334. <https://doi.org/10.1016/j.rse.2005.10.021>
- Caballero, G. R., Delegido, J., Verrelst, J., Pipia, L., Muñoz, J., Burriel, H., Belda, S., Amin, E., Verdú, A. R., & Moreno, J. (2018). ESTIMACIÓN DEL LAI DE LA VEGETACIÓN A PARTIR DE LA SINERGIA SENTINEL 1-SENTINEL 2. In *XVIII Simposio internacional de la sociedad latinoamericana de percepción remota y sistemas de información espacial*. <https://doi.org/10.2307/j.ctv5cg7v4.6>
- Cambell, J. B., & Wynne, R. H. (2011). *Introduction to Remote Sensing* (T. G. Press (ed.)).
- Campos-Taberner, M., García-Haro, F. J., Martínez, B., Sánchez-Ruiz, S., & Gilabert, M. A. (2019). A copernicus sentinel-1 and sentinel-2 classification framework for the 2020+ European common agricultural policy: A case study in València (Spain). *Agronomy*, 9(9). <https://doi.org/10.3390/agronomy9090556>
- CE. (2020). *Copernicus*. <https://www.copernicus.eu/es>
- Chen, J., Deng, F., & Chen, M. (2006). Locally adjusted cubic-spline capping for reconstructing seasonal trajectories of a satellite-derived surface parameter. *IEEE Transactions on Geoscience and Remote Sensing*, 44(8), 2230–2238. <https://doi.org/10.1109/TGRS.2006.872089>
- Chen, J., Jönsson, P., Tamura, M., Gu, Z., Matsushita, B., & Eklundh, L. (2004). A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. *Remote Sensing of Environment*, 91(3–4), 332–344. <https://doi.org/10.1016/j.rse.2004.03.014>
- Chuvieco, E. (1995). *FUNDAMENTOS DE LA TELEDETECCIÓN ESPACIAL* (E. R. S.A. (ed.)). <http://cursosihlla.bdh.org.ar/Sist. Cart. y Teledet./Bibliografia/FUNDAMENTOS-DE-TELEDETECCION-EMILIO-CHUVIECO.pdf>
- Chuvieco, E. (2006). *Teledetección ambiental* (A. Editorial (ed.)).
- Cibula, W. G., Zetka, E. F., & Rickman, D. L. (1992). Response of thematic mapper bands to plant water stress. *International Journal of Remote Sensing*, 13(10), 1869–1880. <https://doi.org/10.1080/01431169208904236>
- Dash, J., Jeganathan, C., & Atkinson, P. M. (2010). The use of MERIS Terrestrial Chlorophyll Index to study spatio-temporal variation in vegetation phenology over India. *Remote Sensing of Environment*, 114(7), 1388–1402. <https://doi.org/10.1016/j.rse.2010.01.021>
- Daughtry, C. S. T., Walthall, C. L., Kim, M. S., De Colstoun, E. B., & McMurtrey, J. E. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sensing of Environment*, 74(2), 229–239. [https://doi.org/10.1016/S0034-4257\(00\)00113-9](https://doi.org/10.1016/S0034-4257(00)00113-9)
- Frampton, W. J., Dash, J., Watmough, G., & Milton, E. J. (2013). Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 82, 83–92. <https://doi.org/10.1016/j.isprsjprs.2013.04.007>

- Gitelson, A. A., Gritz, Y., & Merzlyak, M. N. (2003). Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology*, 160(3), 271–282. <https://doi.org/10.1078/0176-1617-00887>
- Gitelson, A. A., Vina, A., Arkebauer, T. J., Rundquist, D. C., Keydan, G., & Leavitt, B. (2003). Remote estimation of leaf area index and green leaf biomass in maize canopies. *Geophysical Research Letters*, 30(5). <https://doi.org/10.1029/2002gl016450>
- Gitelson, A., & Merzlyak, M. N. (1994). Spectral Reflectance Changes Associated with Autumn Senescence of *Aesculus hippocastanum* L. and *Acer platanoides* L. Leaves. Spectral Features and Relation to Chlorophyll Estimation. *Journal of Plant Physiology*, 143(3), 286–292. [https://doi.org/10.1016/S0176-1617\(11\)81633-0](https://doi.org/10.1016/S0176-1617(11)81633-0)
- Glenn, E. P., Huete, A. R., Nagler, P. L., & Nelson, S. G. (2008). Relationship Between Remotely-sensed Vegetation Indices, Canopy Attributes and Plant Physiological Processes: What Vegetation Indices Can and Cannot Tell Us About the Landscape. *Sensors*, 8, 2136–2160. [www.mdpi.org/sensors](http://www.mdpi.org/sensors)
- Gu, J., Li, X., Huang, C., & Okin, G. S. (2009). A simplified data assimilation method for reconstructing time-series MODIS NDVI data. *Advances in Space Research*, 44(4), 501–509. <https://doi.org/https://doi.org/10.1016/j.asr.2009.05.009>
- Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., & Strachan, I. B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment*, 90(3), 337–352. <https://doi.org/10.1016/j.rse.2003.12.013>
- Huang, Y., CHEN, Z. xin, YU, T., HUANG, X. zhi, & GU, X. fa. (2018). Agricultural remote sensing big data: Management and applications. *Journal of Integrative Agriculture*, 17(9), 1915–1931. [https://doi.org/10.1016/S2095-3119\(17\)61859-8](https://doi.org/10.1016/S2095-3119(17)61859-8)
- Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(1–2), 195–213. [https://doi.org/10.1016/S0034-4257\(02\)00096-2](https://doi.org/10.1016/S0034-4257(02)00096-2)
- Huete, A. R. (1988). A Soil-Adjusted Vegetation Index (SAVI). *Remote Sensing of Environment*, 25(1), 295–309.
- Jönsson, P., & Eklundh, L. (2002). Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Transactions on Geoscience and Remote Sensing*, 40(8), 1824–1832. <https://doi.org/10.1109/TGRS.2002.802519>
- Jordan, C. F. (1969). Derivation of Leaf-Area Index from Quality of Light on the Forest Floor. *Ecology*, 50(4), 663–666. <https://doi.org/10.2307/1936256>
- Joshi, N., Baumann, M., Ehammer, A., Fensholt, R., Grogan, K., Hostert, P., Jepsen, M. R., Kuemmerle, T., Meyfroidt, P., Mitchard, E. T. A., Reiche, J., Ryan, C. M., & Waske, B. (2016). A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring. *Remote Sensing*, 8(1), 1–23. <https://doi.org/10.3390/rs8010070>
- Julien, Y., & Sobrino, J. A. (2010). Comparison of cloud-reconstruction methods for time series of composite NDVI data. *Remote Sensing of Environment*, 114(3), 618–625. <https://doi.org/https://doi.org/10.1016/j.rse.2009.11.001>
- Kandasamy, S., Baret, F., Verger, A., Neveux, P., & Weiss, M. (2013). A comparison of methods for smoothing and gap filling time series of remote sensing observations – application to MODIS LAI products. *Biogeosciences*, 10(6), 4055–4071. <https://doi.org/10.5194/bg-10->

4055-2013

- Kim, Y., & Van Zyl, J. J. (2009). A time-series approach to estimate soil moisture using polarimetric radar data. *IEEE Transactions on Geoscience and Remote Sensing*, 47(8), 2519–2527. <https://doi.org/10.1109/TGRS.2009.2014944>
- Liaghat, S., & Balasundram, S. K. (2010). A review: The role of remote sensing in precision agriculture. *American Journal of Agricultural and Biological Science*, 5(1), 50–55. <https://doi.org/10.3844/ajabssp.2010.50.55>
- Liu, R., Shang, R., Liu, Y., & Lu, X. (2017). Global evaluation of gap-filling approaches for seasonal NDVI with considering vegetation growth trajectory, protection of key point, noise resistance and curve stability. *Remote Sensing of Environment*, 189, 164–179. <https://doi.org/10.1016/j.rse.2016.11.023>
- Lopes, A., Nezry, E., Touzi, R., & Laur, H. (1990). Maximum A Posteriori Speckle Filtering And First Order Texture Models In Sar Images. *10th Annual International Symposium on Geoscience and Remote Sensing*, 2409–2412.
- Lovell, J. L., & Graetz, R. D. (2001). Filtering Pathfinder AVHRR Land NDVI data for Australia. *International Journal of Remote Sensing*, 22(13), 2649–2654. <https://doi.org/10.1080/01431160116874>
- Lu, X., Liu, R., Liu, J., & Liang, S. (2007). Removal of noise by wavelet method to generate high quality temporal data of terrestrial MODIS products. *Photogrammetric Engineering and Remote Sensing*, 73(10), 1129–1139. <https://doi.org/10.14358/PERS.73.10.1129>
- Lymburner, L., Beggs, P. J., & Jacobson, C. R. (2000). Estimation of Canopy-Average Surface-Specific Leaf Area using Landsat TM Data. *Photogrammetric Engineering and Remote Sensing*, 66(2), 183–191.
- Ma, M., & Veroustraete, F. (2006). Reconstructing pathfinder AVHRR land NDVI time-series data for the Northwest of China. *Advances in Space Research*, 37(4), 835–840. <https://doi.org/10.1016/j.asr.2005.08.037>
- Main, R., Cho, M. A., Mathieu, R., O’Kennedy, M. M., Ramoelo, A., & Koch, S. (2011). An investigation into robust spectral indices for leaf chlorophyll estimation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(6), 751–761. <https://doi.org/10.1016/j.isprsjprs.2011.08.001>
- Moreno, Á., García-Haro, F. J., Martínez, B., & Gilabert, M. A. (2014). Noise Reduction and Gap Filling of fAPAR Time Series Using an Adapted Local Regression Filter. *Remote Sensing*, 6(9), 8238–8260. <https://doi.org/10.3390/rs6098238>
- Mulla, D. J. (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering*, 114(4), 358–371. <https://doi.org/10.1016/j.biosystemseng.2012.08.009>
- Nasirzadehdizaji, R., Sanli, F. B., Abdikan, S., Cakir, Z., Sekertekin, A., & Ustuner, M. (2019). Sensitivity analysis of multi-temporal Sentinel-1 SAR parameters to crop height and canopy coverage. *Applied Sciences (Switzerland)*, 9(4). <https://doi.org/10.3390/app9040655>
- Orynbaikyzy, A., Gessner, U., Mack, B., & Conrad, C. (2020). Crop type classification using fusion of sentinel-1 and sentinel-2 data: Assessing the impact of feature selection, optical data availability, and parcel sizes on the accuracies. *Remote Sensing*, 12(17). <https://doi.org/10.3390/RS12172779>
- Periasamy, S. (2018). Significance of dual polarimetric synthetic aperture radar in biomass retrieval: An attempt on Sentinel-1 Keywords: Sentinel 1 Backscattering coefficient Dual-pol scatter plot Dual depolarization index Biomass model. In *Remote Sensing of*

- Environment*. <https://doi.org/10.1016/j.rse.2018.09.003>
- Pipia, L., Muñoz-Marí, J., Amin, E., Belda, S., Camps-Valls, G., & Verrelst, J. (2019). *Fusing optical and SAR time series for LAI gap filling with multioutput Gaussian processes*. <https://doi.org/10.1016/j.rse.2019.111452>
- Poursanidis, D., & Chrysoulakis, N. (2017). Remote Sensing, natural hazards and the contribution of ESA Sentinels missions. *Remote Sensing Applications: Society and Environment*, 6(February), 25–38. <https://doi.org/10.1016/j.rsase.2017.02.001>
- Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., & Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote Sensing of Environment*, 48(2), 119–126. [https://doi.org/10.1016/0034-4257\(94\)90134-1](https://doi.org/10.1016/0034-4257(94)90134-1)
- Roujean, J. L., & Breon, F. M. (1995). Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sensing of Environment*, 51(3), 375–384. [https://doi.org/10.1016/0034-4257\(94\)00114-3](https://doi.org/10.1016/0034-4257(94)00114-3)
- Rouse, J. W., Jr., Haas, R. H., Schell, J. A., & Deering, D. W. (1973). Monitoring vegetation systems in the great plains with ERTS. *ERTS Third Symposium*, 24(1), 309–317. <https://doi.org/10.1021/jf60203a024>
- Small, D. (2011). Flattening Gamma: Radiometric Terrain Correction for SAR Imagery. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, 49(8). <https://doi.org/10.1109/TGRS.2011.2120616>
- Small, D., & Schubert, A. (2008). *Guide to ASAR Geocoding*.
- Sonobe, R., Yamaya, Y., Tani, H., Wang, X., Kobayashi, N., & Mochizuki, K. (2018). Crop classification from Sentinel-2-derived vegetation indices using ensemble learning. *Journal of Applied Remote Sensing*, 12(02), 1. <https://doi.org/10.1117/1.jrs.12.026019>
- Steinhausen, M. J., Wagner, P. D., Narasimhan, B., & Waske, B. (2018). Combining Sentinel-1 and Sentinel-2 data for improved land use and land cover mapping of monsoon regions. *International Journal of Applied Earth Observation and Geoinformation*, 73, 595–604. <https://doi.org/10.1016/j.jag.2018.08.011>
- Trudel, M., Charbonneau, F., & Leconte, R. (2012). Using RADARSAT-2 polarimetric and ENVISAT-ASAR dual-polarization data for estimating soil moisture over agricultural fields. *Canadian Journal of Remote Sensing*, 38(4), 514–527. <https://doi.org/10.5589/m12-043>
- Van Tricht, K., Gobin, A., Gilliams, S., & Piccard, I. (2018). Synergistic use of radar sentinel-1 and optical sentinel-2 imagery for crop mapping: A case study for Belgium. *Remote Sensing*, 10(10). <https://doi.org/10.3390/rs10101642>
- Veloso, A., Mermoz, S., Bouvet, A., Le Toan, T., Planells, M., Dejoux, J. F., & Ceschia, E. (2017). Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. *Remote Sensing of Environment*, 199, 415–426. <https://doi.org/10.1016/j.rse.2017.07.015>
- Vivoy, N., Arino, O., & Belward, A. S. (1992). The Best Index Slope Extraction (BISE): A method for reducing noise in NDVI time-series. *International Journal of Remote Sensing*, 13(8), 1585–1590. <https://doi.org/10.1080/01431169208904212>
- Wilson, E. H., & Sader, S. A. (2002). Detection of forest harvest type using multiple dates of Landsat TM imagery. *Remote Sensing of Environment*, 80(3), 385–396. [https://doi.org/10.1016/S0034-4257\(01\)00318-2](https://doi.org/10.1016/S0034-4257(01)00318-2)
- Zhang, Z., & Ke, C. (2016). Monitoring and analysis of changes in a wetland landscape in Xingzi

county. *Earth Science Informatics*, 9(1), 35–45. <https://doi.org/10.1007/s12145-015-0232-4>

Zhu, W., Pan, Y., He, H., Wang, L., Mou, M., & Liu, J. (2012). A Changing-Weight Filter Method for Reconstructing a High-Quality NDVI Time Series to Preserve the Integrity of Vegetation Phenology. *IEEE Transactions on Geoscience and Remote Sensing*, 50(4), 1085–1094. <https://doi.org/10.1109/TGRS.2011.2166965>



# ANEXOS

## ÍNDICE DE ANEXOS

ANEXO I. INFORMACIÓN DE LOS 95 CULTIVOS DE PARTIDA.....	70
ANEXO II. MATRICES DE CORRELACIÓN.....	73
ANEXO III. COMPORTAMIENTO REFLEXIVO Y RETRODISPERSIVO DE LOS CULTIVOS .....	82
ANEXO IV. RMSE DE EVALUACIÓN DE IMPACTO DE HUECOS .....	88
ANEXO V. SERIES DE NDVI DE REFERENCIA PARA EL RELLENO DE HUECOS A PARTIR DE DATOS DE SENTINEL-1 .....	91
ANEXO VI. RMSE DE LAS TÉCNICAS DE RELLENO DE HUECOS .....	93
ANEXO VII. SCRIPTS DE PYTHON .....	100

## ANEXO I. INFORMACIÓN DE LOS 95 CULTIVOS DE PARTIDA

CULTIVO	nº parcelas	Superficie (ha)
CEBADA	4157	9354,46
TRIGO BLANDO	3472	6844,76
PASTO ARBUSTIVO DE 5 O MAS AÑOS	3398	4952,71
BARBECHO TRADICIONAL	2896	3584,16
VIÑEDO VINIFICACIÓN	2787	3508,60
PASTOS PERMANENTES DE 5 O MAS AÑOS	2373	2942,67
MAÍZ	1570	4870,64
BARBECHO SIN PRODUCCIÓN	1091	1870,21
OLIVAR	797	588,51
ALFALFA	519	935,92
OTRAS SUPERFICIES FORESTALES	370	894,20
Otras utilizaciones no agrarias ni forestales	273	80,28
HUERTA	270	54,43
ARROZ	264	492,97
RAYGRAS PERENNE	259	153,47
GIRASOL	253	860,96
ALMENDROS	243	295,11
GUISANTES	228	640,60
FESTUCA	198	114,58
PASTIZAL DE 5 O MAS AÑOS	185	105,56
COLZA	180	601,17
VEZA	176	307,54
TOMATE PARA TRANSFORMACIÓN	174	649,64
AVENA	168	624,97
CULTIVOS MIXTOS DE ESPECIES PRATENSES	121	86,60
ESPÁRRAGOS	104	106,20
RAYGRAS ANUAL ( <i>Lolium multiflorum</i> LAM. And hybrids)	103	143,43
OTROS FRUTALES	100	12,12
MAÍZ DULCE	84	272,17
LECHUGA	80	97,92
PIMIENTO	68	93,28
ALCACHOFA	59	35,41
RASTROJERAS	57	62,22
PERALES	53	16,19
TRITICALE	52	126,54
PASTOS DE MENOS DE 5 AÑOS	49	12,90
PASTO ARBOLADO DE 5 O MAS AÑOS	44	97,53
REMOLACHA	41	146,55
HABAS	38	104,07
CEBOLLA	36	142,17
FORESTACIONES VINCULADAS AL REGLAMENTO Nº 1698/2005	32	123,47

CULTIVO	nº parcelas	Superficie (ha)
TOMATE	30	21,32
BARBECHO MEDIOAMBIENTAL	29	41,00
VIÑA - OLIVAR	25	22,91
MANZANOS	24	20,34
MENTA	24	8,46
ENDIVIA	22	56,03
PATATA	22	28,43
ESPINACA	21	109,00
AJO	19	6,74
BRÓCOLI	18	62,74
ENDRINO o ARAÑÓN	15	19,43
COLIFLOR	14	6,06
MELOCOTONEROS	14	7,51
TRIGO DURO	14	19,46
CALABAZA	13	15,02
CEREZOS	13	18,83
GARBANZOS	12	15,55
MAIZ	12	88,93
NOGALES	12	24,94
PORTAINJERTOS DE VID	12	52,42
CARDO	11	23,66
CHOPOS	10	3,84
TITARROS	10	15,28
VIVEROS	9	7,33
JUDÍA	8	16,44
MEZCLA VEZA - AVENA	8	11,93
PASTOS HERBACEOS	8	65,35
ACELGA	7	5,99
CALABACÍN	7	2,15
MEZCLA GUISANTE - AVENA	7	37,88
ALUBIAS	5	34,56
BARBECHO	5	42,81
ESCAROLA	5	37,89
FLORES	5	1,39
LAVANDA	5	24,72
LÚPULO	5	8,47
MEZCLA AVENA-CEBADA	5	2,04
PISTACHO	5	16,27
TRIGO	5	66,29
FRUTOS DE CÁSCARA	4	2,17
PASTOS ARBUSTIVOS	4	59,62
CROTALARIA JUNCEA	3	8,21
MEMBRILLO	3	2,95
TRITICUM SPELTA	3	38,11

CULTIVO	nº parcelas	Superficie (ha)
HIERBABUENA	2	0,72
LAVANDÍN	2	4,60
CENTENO	1	7,00
ESPARCETA	1	1,37
FRUTALES	1	11,73
GUISANTE	1	5,46
MISCANTHUS	1	2,06
OTRAS CRUCIFERAS	1	5,04
PASTO DEL SUDÁN	1	1,11
SUPERFICIES FORESTALES MADERABLES	1	1,43

## ANEXO II. MATRICES DE CORRELACIÓN

### MATRICES DE CORRELACIÓN ENTRE ÍNDICES DE SENTINEL-1 Y SENTINEL-2 SIN AGRUPACIONES

IDPDD	0.55	0.38	0.56	0.49	0.42	0.68	0.55	0.41	0.51	0.52
RDIF	-0.55	-0.38	-0.56	-0.48	-0.42	-0.68	-0.55	-0.41	-0.51	-0.52
RDIV	0.59	0.52	0.59	0.65	0.52	0.76	0.6	0.51	0.55	0.61
RMULTI	-0.4	-0.26	-0.4	-0.28	-0.31	-0.41	-0.39	-0.31	-0.38	-0.35
RSUM	-0.51	-0.33	-0.51	-0.41	-0.39	-0.6	-0.5	-0.38	-0.47	-0.47
RVI	0.59	0.51	0.59	0.65	0.51	0.76	0.59	0.5	0.55	0.61
VDDPI	0.59	0.52	0.59	0.65	0.52	0.76	0.6	0.51	0.55	0.61
VH	-0.19	-0.064	-0.17	-0.019	-0.12	-0.1	-0.14	-0.11	-0.16	-0.13
WV	-0.53	-0.36	-0.54	-0.45	-0.41	-0.64	-0.53	-0.4	-0.49	-0.49
	CI-RE	MCARI	MTCI	NDMI	NDVI	NDVire	RDVI	SAVI	SeLI	SLAVI

**Figura 47.** Matriz de correlaciones medias de las parcelas del cultivo de cebada entre los índices Sentinel-1 (pasada DESC\_81) y Sentinel-2.

IDPDD	-0.39	-0.57	-0.61	-0.58	-0.52	-0.17	-0.56	-0.53	-0.5	-0.43
RDIF	0.39	0.57	0.61	0.57	0.52	0.17	0.55	0.53	0.49	0.43
RDIV	0.29	0.14	0.083	0.18	0.14	0.49	0.11	0.12	0.14	0.26
RMULTI	0.55	0.69	0.73	0.7	0.65	0.38	0.68	0.65	0.63	0.59
RSUM	0.44	0.6	0.64	0.62	0.56	0.23	0.59	0.56	0.53	0.48
RVI	0.29	0.15	0.093	0.19	0.16	0.49	0.12	0.14	0.15	0.26
VDDPI	0.29	0.14	0.083	0.18	0.14	0.49	0.11	0.12	0.14	0.26
VH	0.55	0.66	0.66	0.69	0.62	0.43	0.62	0.61	0.58	0.58
WV	0.42	0.59	0.62	0.6	0.54	0.21	0.57	0.55	0.51	0.46
	CI-RE	MCARI	MTCI	NDMI	NDVI	NDVire	RDVI	SAVI	SeLI	SLAVI

**Figura 48.** Matriz de correlaciones medias de las parcelas del cultivo de girasol entre los índices Sentinel-1 (pasada DESC\_81) y Sentinel-2.



IDPDD	0.4	0.48	0.44	0.43	0.47	0.43	0.46	0.47	0.46	0.42
RDIF	-0.4	-0.49	-0.44	-0.43	-0.47	-0.44	-0.46	-0.47	-0.46	-0.42
RDIV	0.66	0.66	0.66	0.7	0.7	0.75	0.66	0.7	0.67	0.64
RMULTI	0.04	-0.075	0.0068	0.025	-0.0079	0.078	-0.033	-0.01	-0.018	-0.011
RSUM	-0.2	-0.3	-0.24	-0.22	-0.26	-0.2	-0.27	-0.26	-0.26	-0.24
RVI	0.65	0.65	0.66	0.69	0.7	0.75	0.65	0.7	0.67	0.63
VDDPI	0.66	0.66	0.66	0.7	0.7	0.75	0.66	0.7	0.67	0.64
VH	0.38	0.28	0.35	0.37	0.37	0.45	0.32	0.37	0.34	0.32
VW	-0.3	-0.39	-0.34	-0.32	-0.36	-0.31	-0.36	-0.36	-0.36	-0.33
	CI-RE	MCARI	MTCI	NDMI	NDVI	NDVire	RDVI	SAVI	SeLI	SLAVI

**Figura 49.** Matriz de correlaciones medias de las parcelas del cultivo de maíz entre los índices Sentinel-1 (pasada DESC\_81) y Sentinel-2.

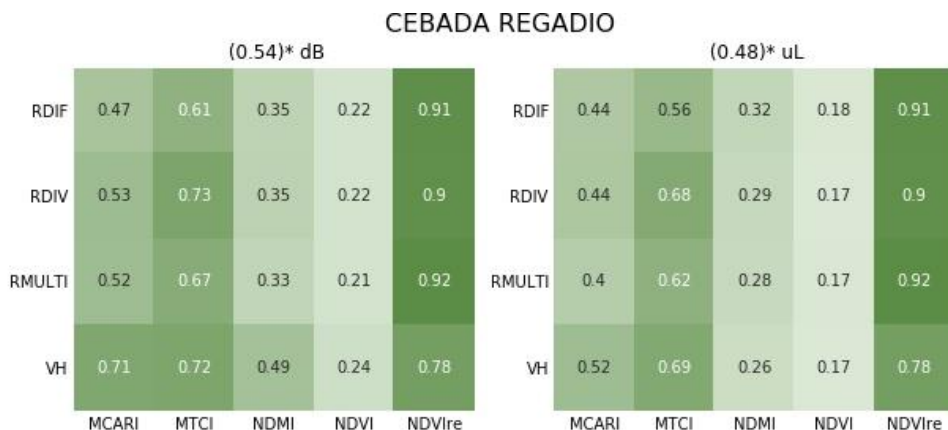
IDPDD	0.59	0.36	0.58	0.54	0.48	0.71	0.58	0.47	0.55	0.56
RDIF	-0.59	-0.35	-0.58	-0.54	-0.47	-0.71	-0.58	-0.46	-0.55	-0.56
RDIV	0.72	0.56	0.68	0.7	0.62	0.86	0.68	0.61	0.66	0.73
RMULTI	-0.48	-0.31	-0.49	-0.45	-0.42	-0.53	-0.5	-0.41	-0.48	-0.46
RSUM	-0.56	-0.33	-0.55	-0.5	-0.45	-0.66	-0.56	-0.44	-0.53	-0.53
RVI	0.71	0.55	0.67	0.69	0.61	0.85	0.67	0.6	0.65	0.71
VDDPI	0.72	0.56	0.68	0.7	0.62	0.86	0.68	0.61	0.66	0.73
VH	-0.3	-0.14	-0.31	-0.23	-0.25	-0.28	-0.3	-0.24	-0.3	-0.26
VW	-0.57	-0.34	-0.56	-0.52	-0.46	-0.68	-0.57	-0.45	-0.54	-0.54
	CI-RE	MCARI	MTCI	NDMI	NDVI	NDVire	RDVI	SAVI	SeLI	SLAVI

**Figura 50.** Matriz de correlaciones medias de las parcelas del cultivo del trigo entre los índices Sentinel-1 (pasada DESC\_81) y Sentinel-2.

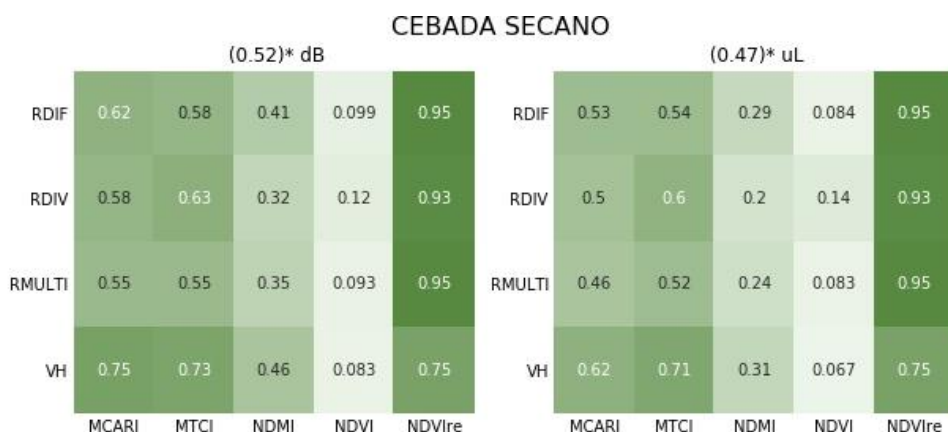
IDPDD	0.29	-0.14	0.41	-0.25	-0.032	0.23	0.23	-0.034	0.21	-0.17
RDIF	-0.29	0.14	-0.42	0.25	0.031	-0.23	-0.23	0.033	-0.21	0.17
RDIV	0.69	0.57	0.76	0.39	0.33	0.7	0.57	0.31	0.42	0.35
RMULTI	-0.16	0.27	-0.24	0.33	0.059	-0.089	-0.12	0.057	-0.16	0.23
RSUM	-0.23	0.21	-0.34	0.31	0.064	-0.16	-0.18	0.064	-0.18	0.22
RVI	0.69	0.57	0.76	0.39	0.33	0.7	0.57	0.32	0.42	0.35
VDDPI	0.69	0.57	0.76	0.39	0.33	0.7	0.57	0.31	0.42	0.35
VH	0.026	0.42	-0.058	0.45	0.17	0.1	0.025	0.16	-0.033	0.34
VW	-0.26	0.18	-0.38	0.28	0.049	-0.19	-0.2	0.051	-0.19	0.2
	CI-RE	MCARI	MTCI	NDMI	NDVI	NDVire	RDVI	SAVI	SeLI	SLAVI

**Figura 51.** Matriz de correlaciones medias de las parcelas del cultivo de la vid entre los índices Sentinel-1 (pasada DESC\_81) y Sentinel-2.

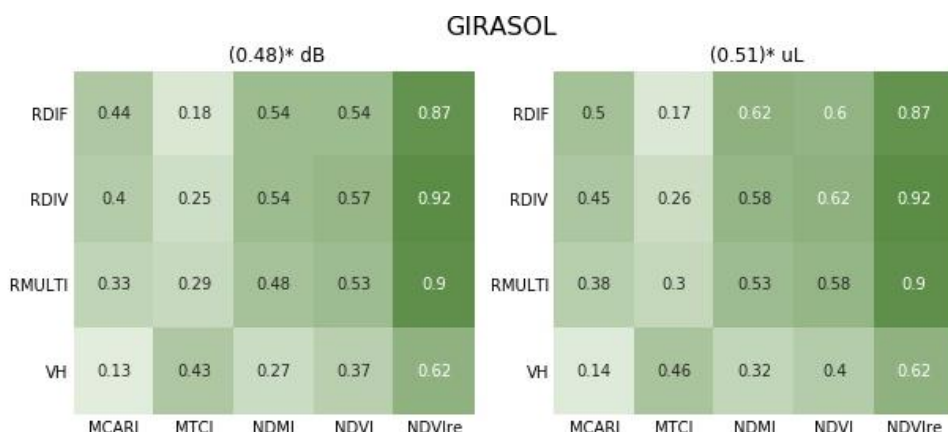
## MATRICES DE CORRELACIÓN ENTRE ÍNDICES DE SENTINEL-1 Y SENTINEL-2 AGRUPADOS, PARA LA COMPARACIÓN DE UNIDADES



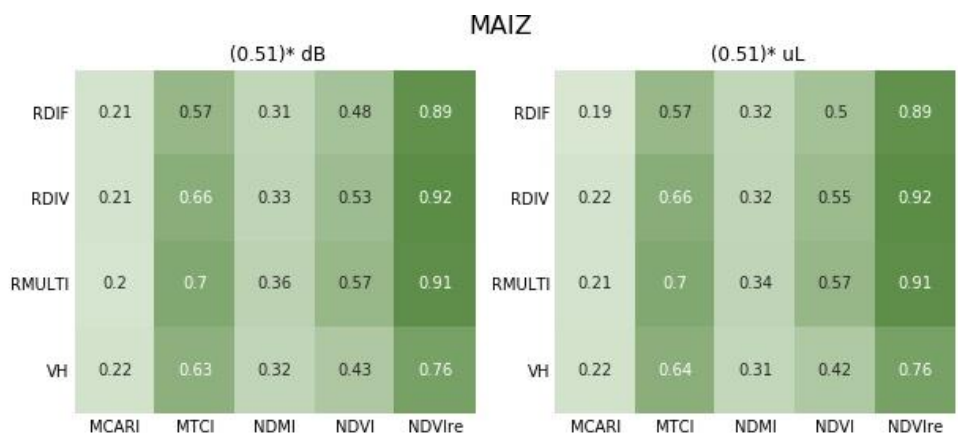
**Figura 52.** Matrices de correlaciones medias (en valor absoluto) de las 75 parcelas del CE de la cebada en regadío entre los índices de S1 y S2 en unidades de decibelios (dB, izquierda) y lineales ( $m^2m^{-2}$ , derecha).



**Figura 53.** Matrices de correlaciones medias (en valor absoluto) de las 75 parcelas del CE de la cebada en secano entre los índices de S1 y S2 en unidades de decibelios (dB, izquierda) y lineales ( $m^2m^{-2}$ , derecha).



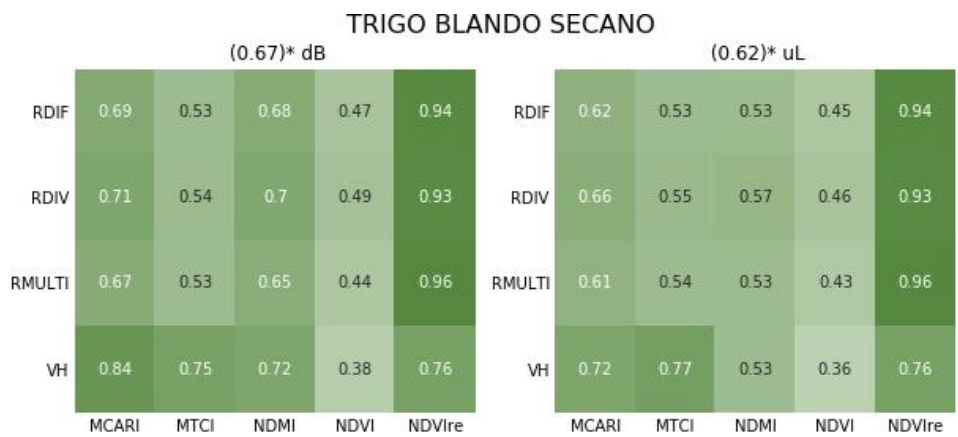
**Figura 54.** Matrices de correlaciones medias (en valor absoluto) de las 75 parcelas del CE del girasol entre los índices de S1 y S2 en unidades de decibelios (dB, izquierda) y lineales ( $m^2m^{-2}$ , derecha)



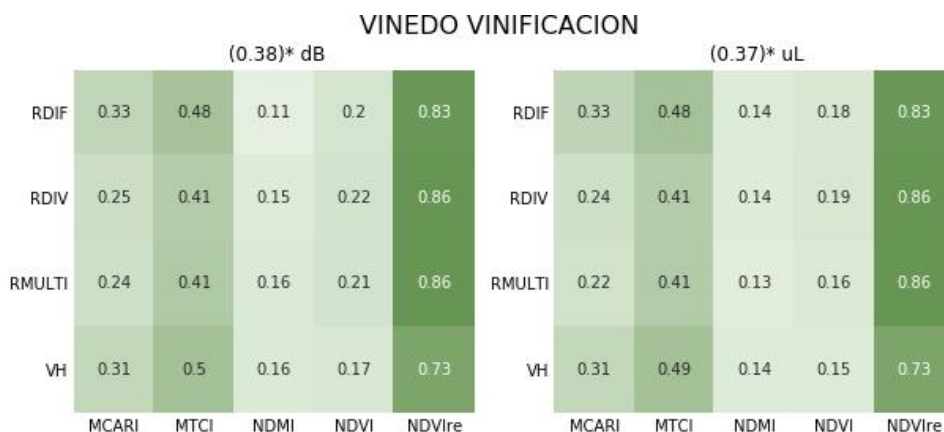
**Figura 55.** Matrices de correlaciones medias (en valor absoluto) de las 75 parcelas del CE del maíz entre los índices de S1 y S2 en unidades de decibelios (dB, izquierda) y lineales (m2m-2, derecha)



**Figura 56.** Matrices de correlaciones medias (en valor absoluto) de las 75 parcelas del CE del trigo blando en regadío entre los índices de S1 y S2 en unidades de decibelios (dB, izquierda) y lineales (m2m-2, derecha)

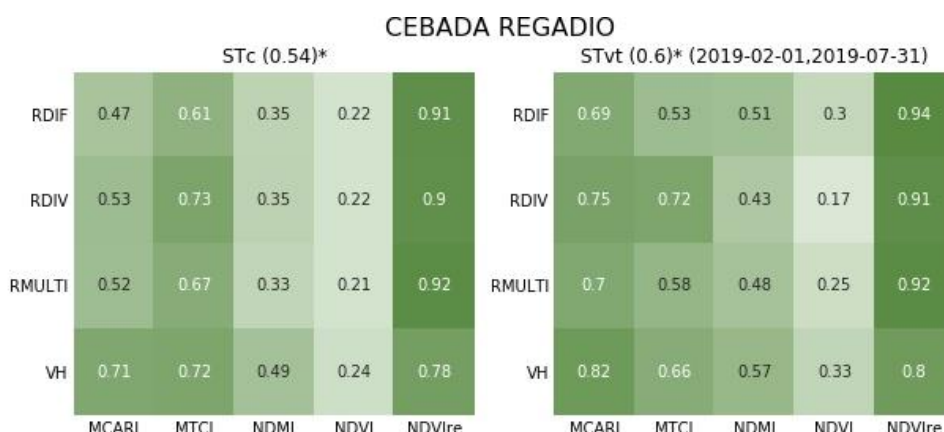


**Figura 57.** Matrices de correlaciones medias (en valor absoluto) de las 75 parcelas del CE del trigo blando en secano entre los índices de S1 y S2 en unidades de decibelios (dB, izquierda) y lineales (m2m-2, derecha)

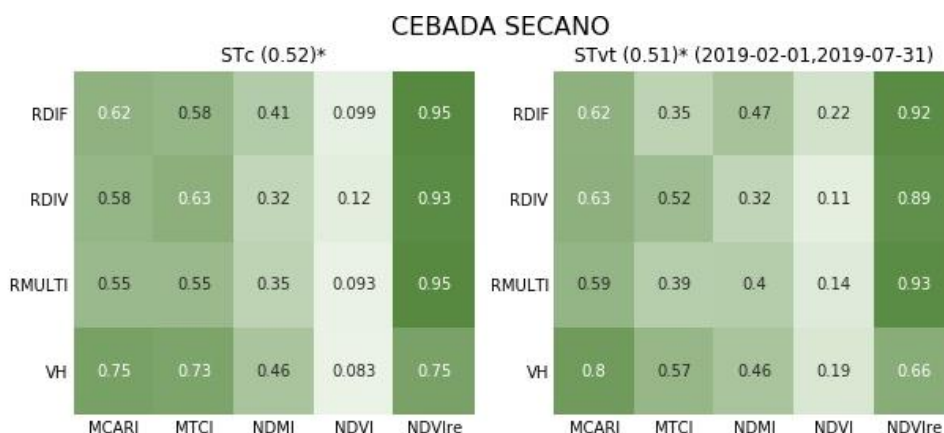


**Figura 58.** Matrices de correlaciones medias (en valor absoluto) de las 75 parcelas del CE de la vid entre los índices de S1 y S2 en unidades de decibelios (dB, izquierda) y lineales (m2m-2, derecha)

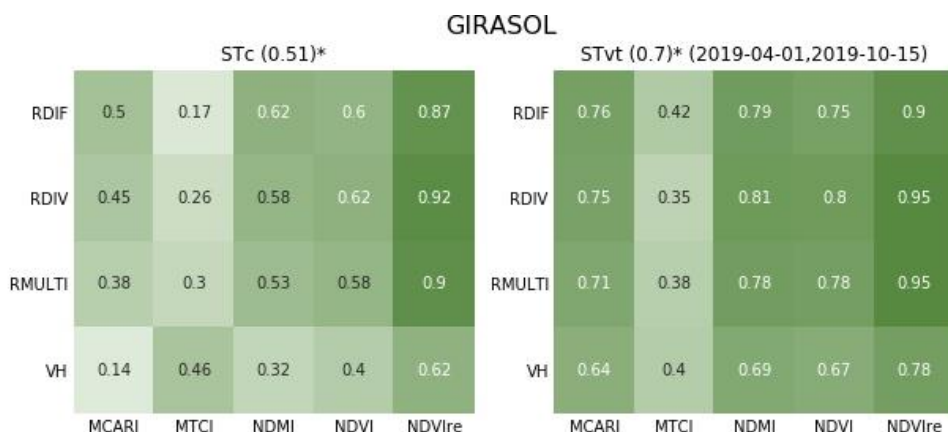
**MATRICES DE CORRELACIÓN ENTRE ÍNDICES DE SENTINEL-1 Y SENTINEL-2  
AGRUPADOS, PARA LA COMPARACIÓN DE VENTANAS TEMPORALES**



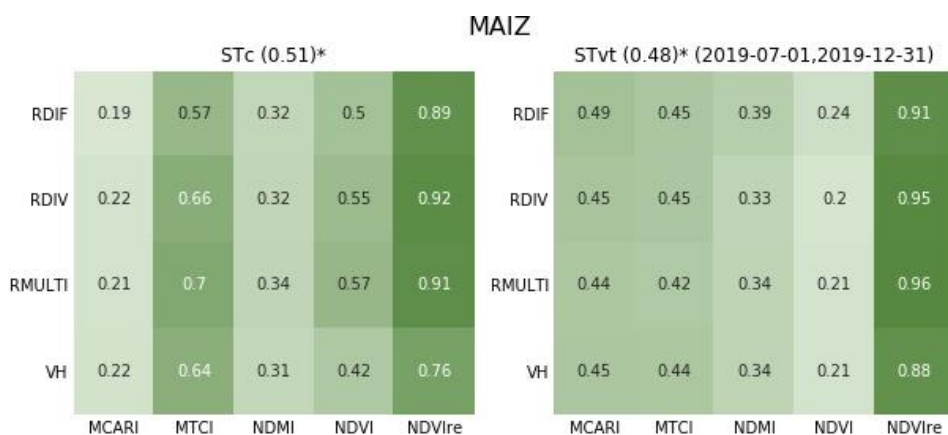
**Figura 59.** Matrices de correlaciones medias (en valor absoluto) de las 75 parcelas del CE de la cebada en regadío entre los índices de Sentinel-1 y Sentinel-2, con datos de toda la serie temporal (ST<sub>C</sub>, izquierda) y calendario fenológico de la cebada (ST<sub>VT</sub>, derecha). \*Media de la matriz de correlaciones medias.



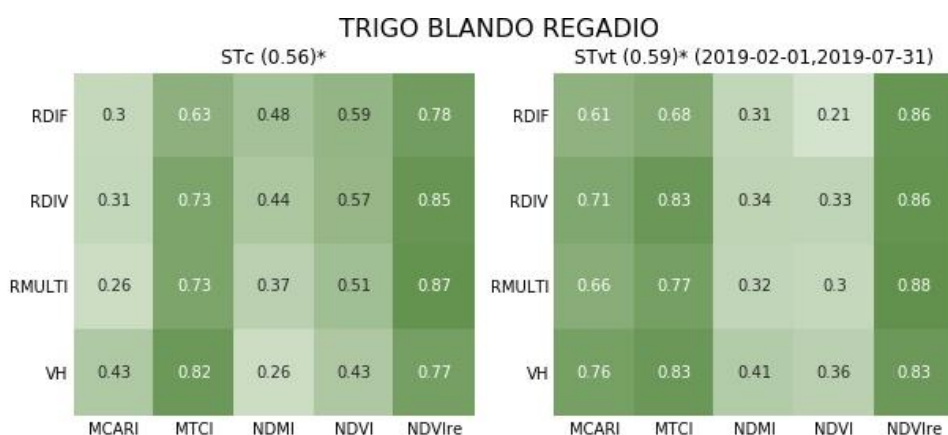
**Figura 60.** Matrices de correlaciones medias (en valor absoluto) de las 75 parcelas del CE de la cebada en secano entre los índices de Sentinel-1 y Sentinel-2, con datos de toda la serie temporal (ST<sub>C</sub>, izquierda) y calendario fenológico de la cebada (ST<sub>VT</sub>, derecha). \*Media de la matriz de correlaciones medias.



**Figura 61.** Matrices de correlaciones medias (en valor absoluto) de las 75 parcelas del CE del girasol entre los índices de Sentinel-1 y Sentinel-2, con datos de toda la serie temporal (ST<sub>C</sub>, izquierda) y calendario fenológico de la cebada (ST<sub>VT</sub>, derecha). \*Media de la matriz de correlaciones medias.

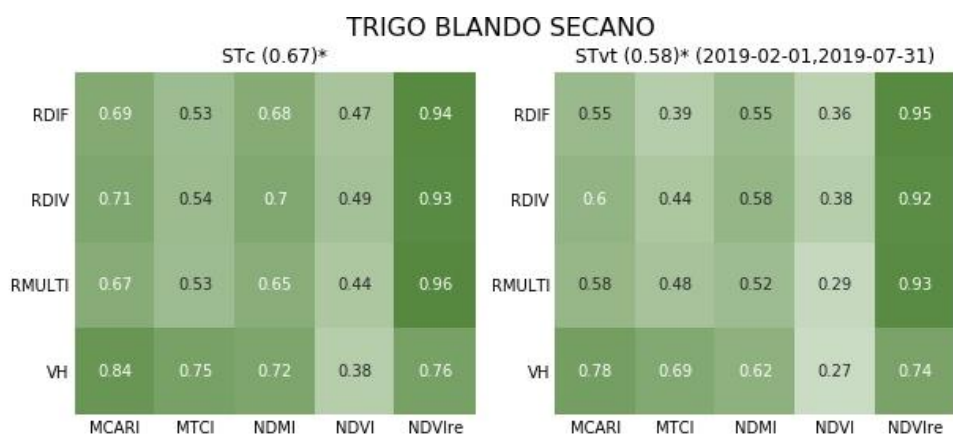


**Figura 62.** Matrices de correlaciones medias (en valor absoluto) de las 75 parcelas del CE de maíz entre los índices de Sentinel-1 y Sentinel-2, con datos de toda la serie temporal (ST<sub>C</sub>, izquierda) y calendario fenológico de la cebada (ST<sub>VT</sub>, derecha). \*Media de la matriz de correlaciones medias.



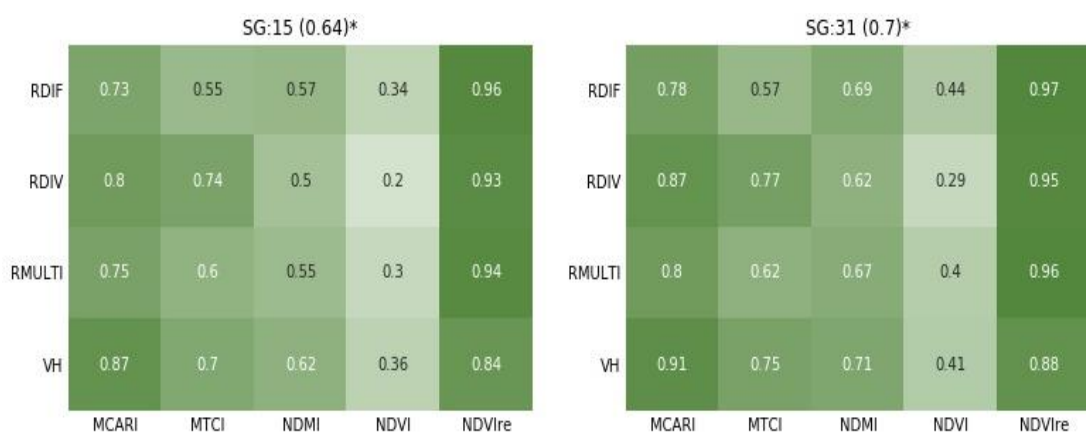
**Figura 63.** Matrices de correlaciones medias (en valor absoluto) de las 75 parcelas del CE del trigo en regadío entre los índices de Sentinel-1 y Sentinel-2, con datos de toda la serie temporal (ST<sub>C</sub>, izquierda) y calendario fenológico de la cebada (ST<sub>VT</sub>, derecha). \*Media de la matriz de correlaciones medias.



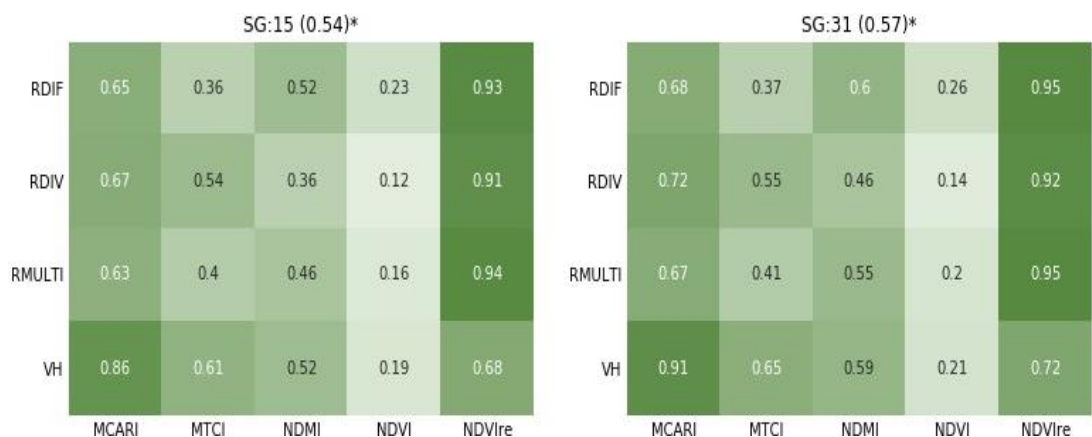


**Figura 64.** Matrices de correlaciones medias (en valor absoluto) de las 75 parcelas del CE del trigo en secano entre los índices de Sentinel-1 y Sentinel-2, con datos de toda la serie temporal (ST<sub>c</sub>, izquierda) y calendario fenológico de la cebada (ST<sub>vt</sub>, derecha). \*Media de la matriz de correlaciones medias.

**MATRICES DE CORRELACIÓN ENTRE ÍNDICES DE SENTINEL-1 Y SENTINEL-2  
AGRUPADOS, PARA LA COMPARACIÓN DEL SUAVIZADO DE CURVAS**

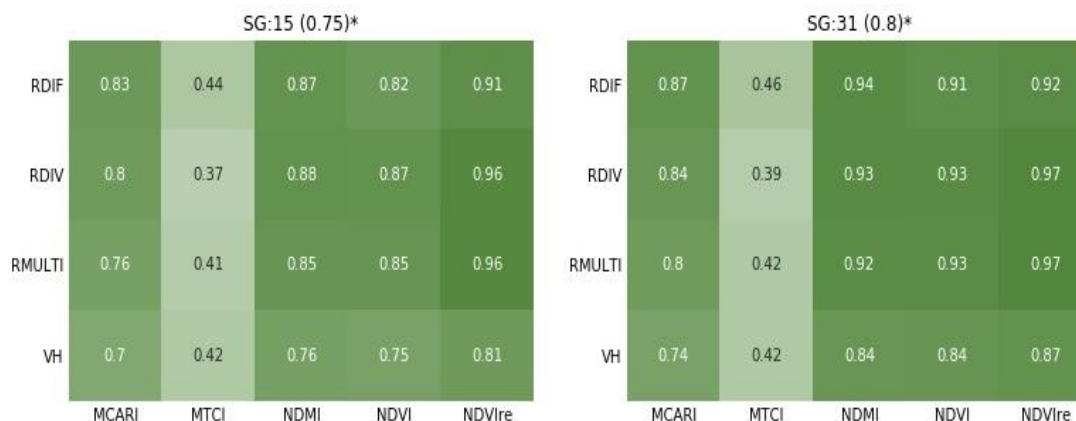


**Figura 65.** Matrices de correlaciones medias de las 75 parcelas del CE de la cebada en regadío entre los índices de Sentinel-1 y Sentinel-2, con filtro SG aplicado para diferente tamaño de ventana: 15 (ST<sub>vt,15</sub>, izquierda) y 31 (ST<sub>vt,31</sub>, derecha). \*Media de la matriz de correlaciones medias.



**Figura 66.** Matrices de correlaciones medias de las 75 parcelas del CE de la cebada en secano entre los índices de Sentinel-1 y Sentinel-2, con filtro SG aplicado para diferente tamaño de ventana: 15 (ST<sub>vt,15</sub>, izquierda) y 31 (ST<sub>vt,31</sub>, derecha). \*Media de la matriz de correlaciones medias.





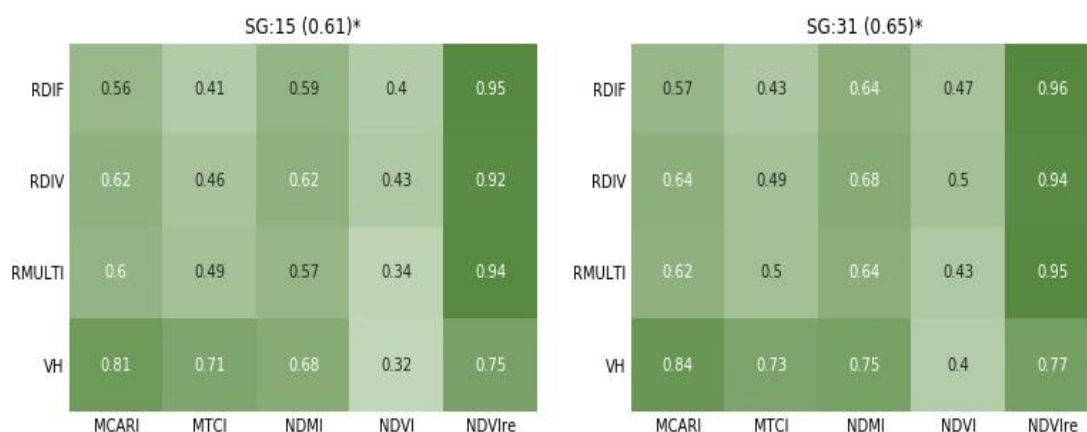
**Figura 67.** Matrices de correlaciones medias de las 75 parcelas del CE del girasol entre los índices de Sentinel-1 y Sentinel-2, con filtro SG aplicado para diferente tamaño de ventana: 15 ( $ST_{VT,15}$ , izquierda) y 31 ( $ST_{VT,31}$ , derecha). \*Media de la matriz de correlaciones medias.



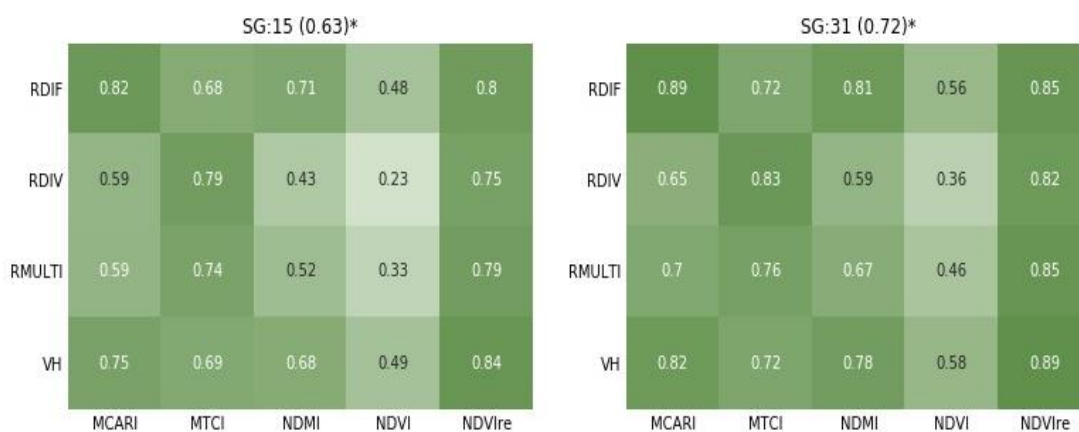
**Figura 68.** Matrices de correlaciones medias de las 75 parcelas del CE del maíz entre los índices de Sentinel-1 y Sentinel-2, con filtro SG aplicado para diferente tamaño de ventana: 15 ( $ST_{VT,15}$ , izquierda) y 31 ( $ST_{VT,31}$ , derecha). \*Media de la matriz de correlaciones medias.



**Figura 69.** Matrices de correlaciones medias de las 75 parcelas del CE del trigo en regadío entre los índices de Sentinel-1 y Sentinel-2, con filtro SG aplicado para diferente tamaño de ventana: 15 ( $ST_{VT,15}$ , izquierda) y 31 ( $ST_{VT,31}$ , derecha). \*Media de la matriz de correlaciones medias.

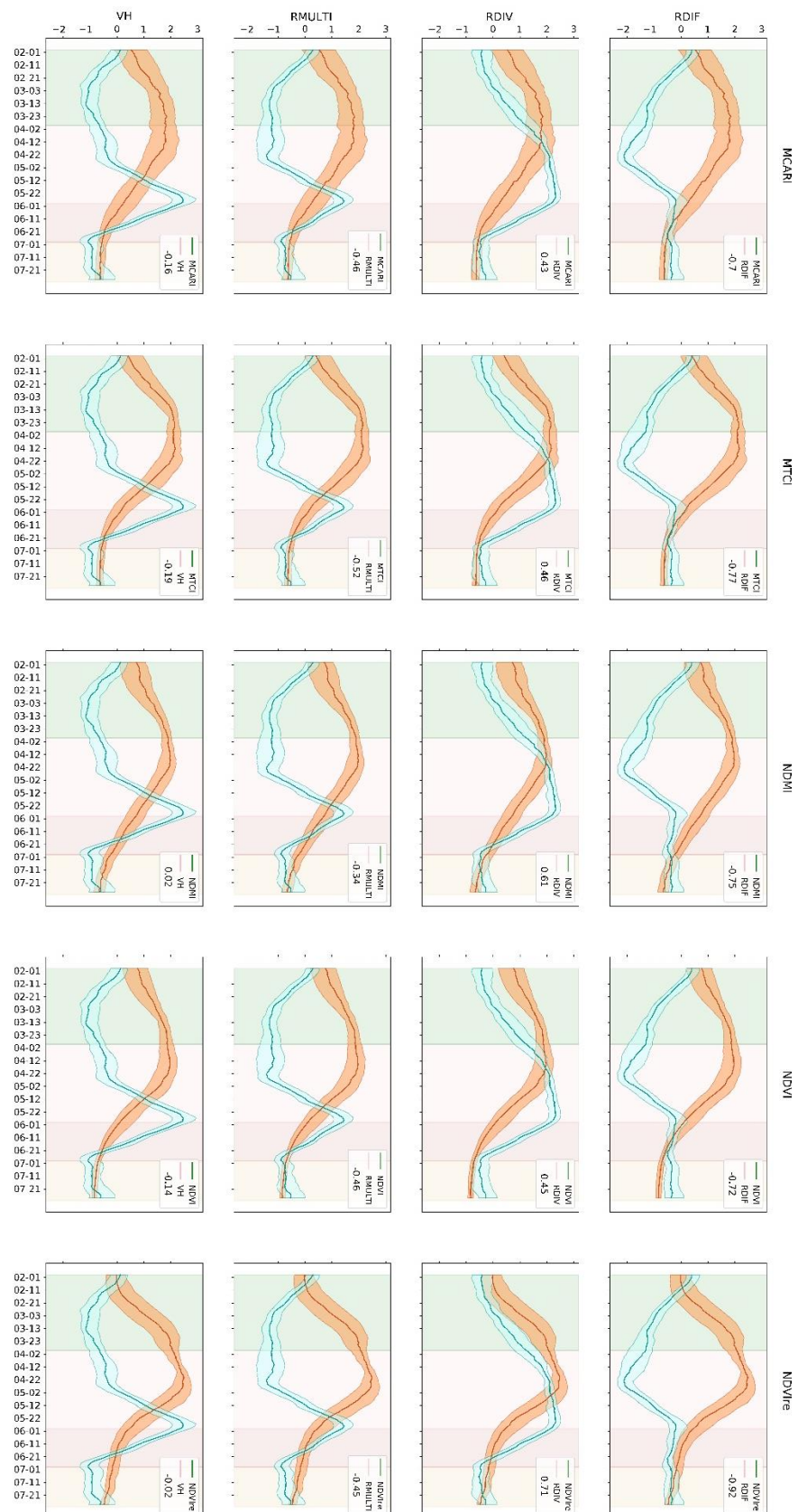


**Figura 70.** Matrices de correlaciones medias de las 75 parcelas del CE del trigo en seco entre los índices de Sentinel-1 y Sentinel-2, con filtro SG aplicado para diferente tamaño de ventana: 15 ( $ST_{VT,15}$ , izquierda) y 31 ( $ST_{VT,31}$ , derecha). \*Media de la matriz de correlaciones medias.

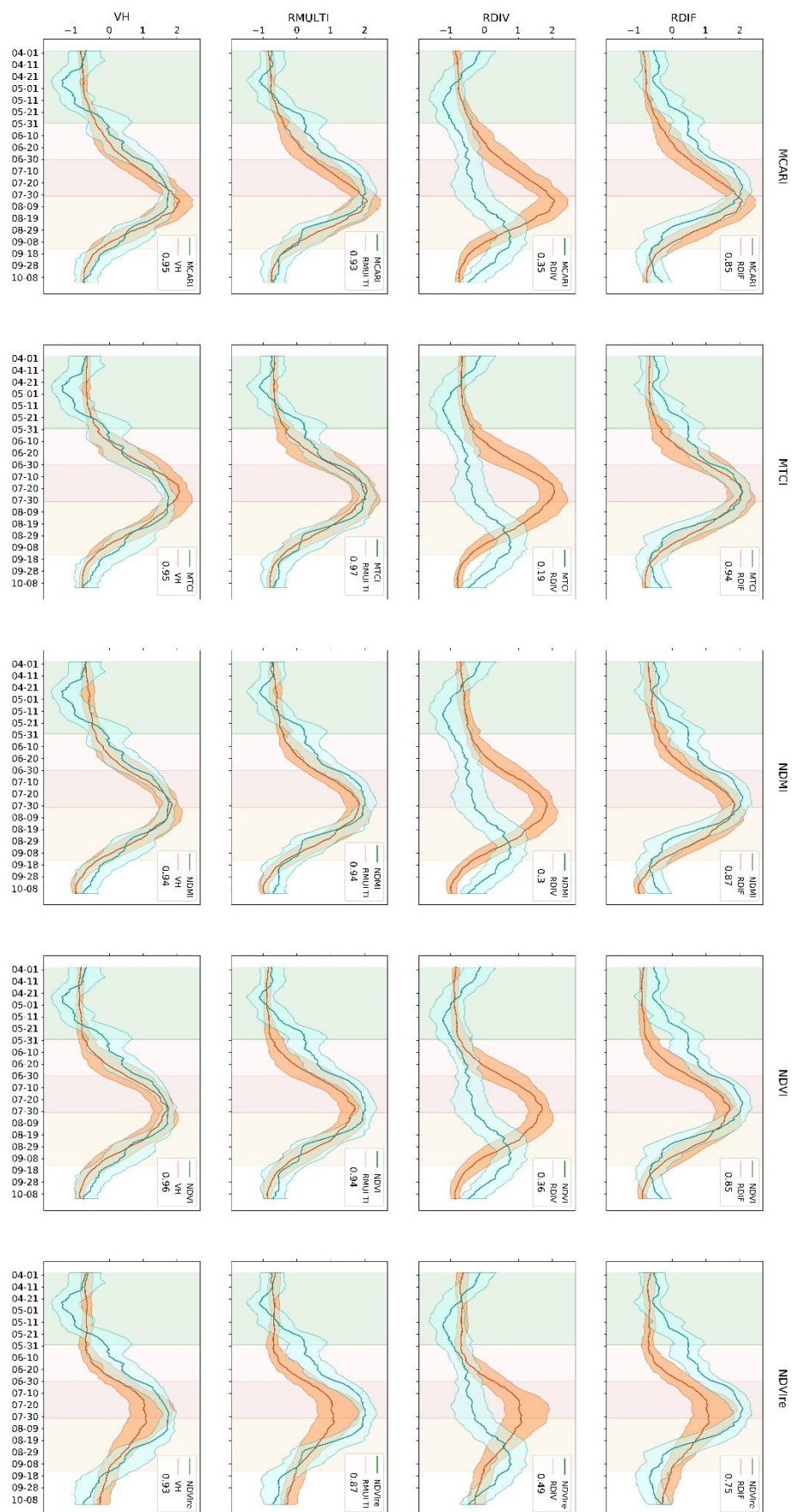


**Figura 71.** Matrices de correlaciones medias de las 75 parcelas del CE de la vid entre los índices de Sentinel-1 y Sentinel-2, con filtro SG aplicado para diferente tamaño de ventana: 15 ( $ST_{VT,15}$ , izquierda) y 31 ( $ST_{VT,31}$ , derecha). \*Media de la matriz de correlaciones medias.

## ANEXO III. COMPORTAMIENTO REFLEXIVO Y RETRODISPERSIVO DE LOS CULTIVOS

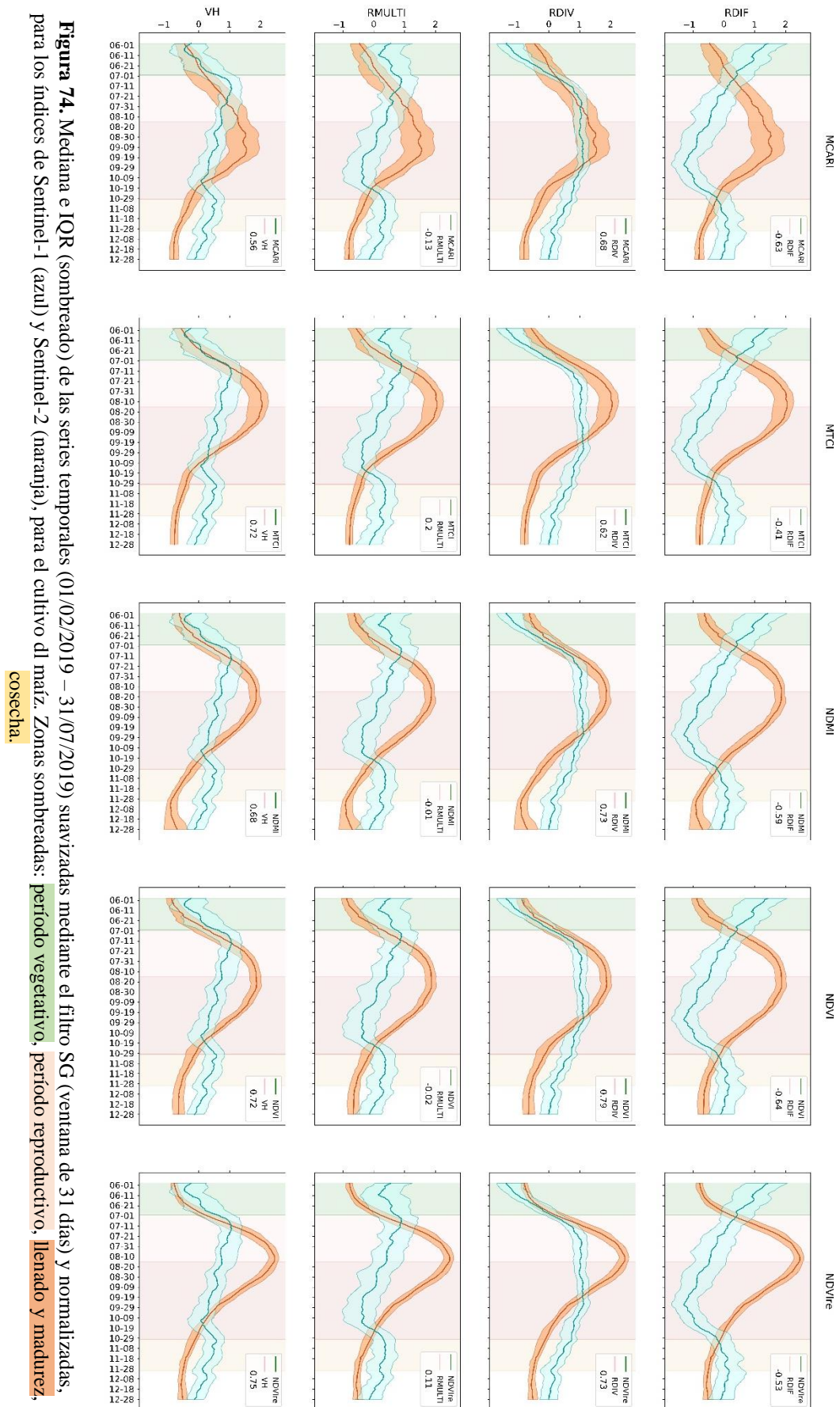


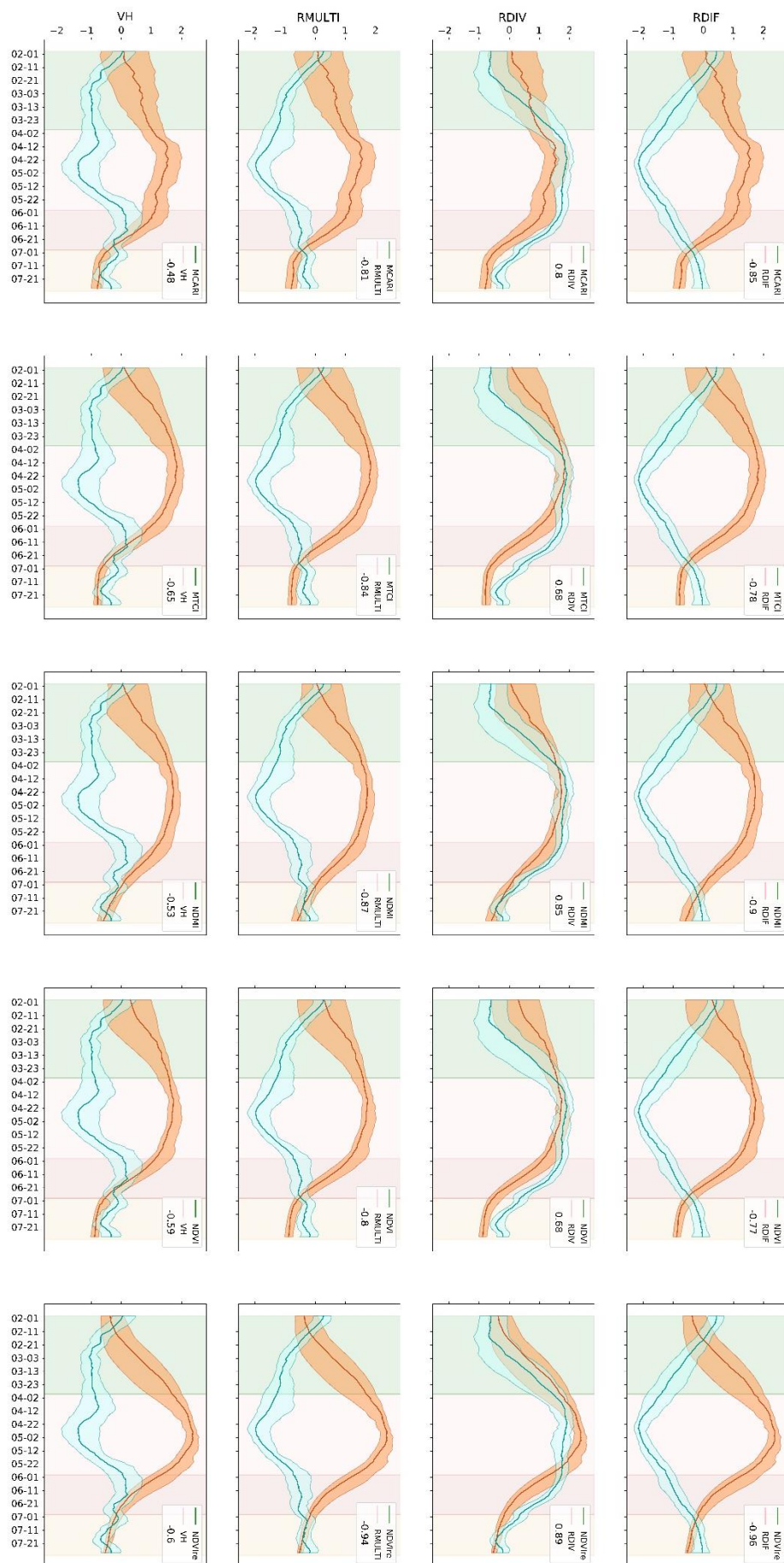
**Figura 72.** Mediana e IQR (sombreado) de las series temporales (01/02/2019 – 31/07/2019) suavizadas mediante el filtro SG (ventana de 31 días) y normalizadas, para los índices de Sentinel-1 (azul) y Sentinel-2 (naranja), para el cultivo de cebada en secano. Zonas sombreadas: período vegetativo, período reproductivo, llenado y madurez, cosecha.



**Figura 73.** Mediana e IQR (sombreado) de las series temporales (01/02/2019 – 31/07/2019) suavizadas mediante el filtro SG (ventana de 31 días) y normalizadas, para los índices de Sentinel-1 (azul) y Sentinel-2 (naranja), para el cultivo del girasol. Zonas sombreadas: período vegetativo, período reproductivo, llenado y madurez, cosecha.

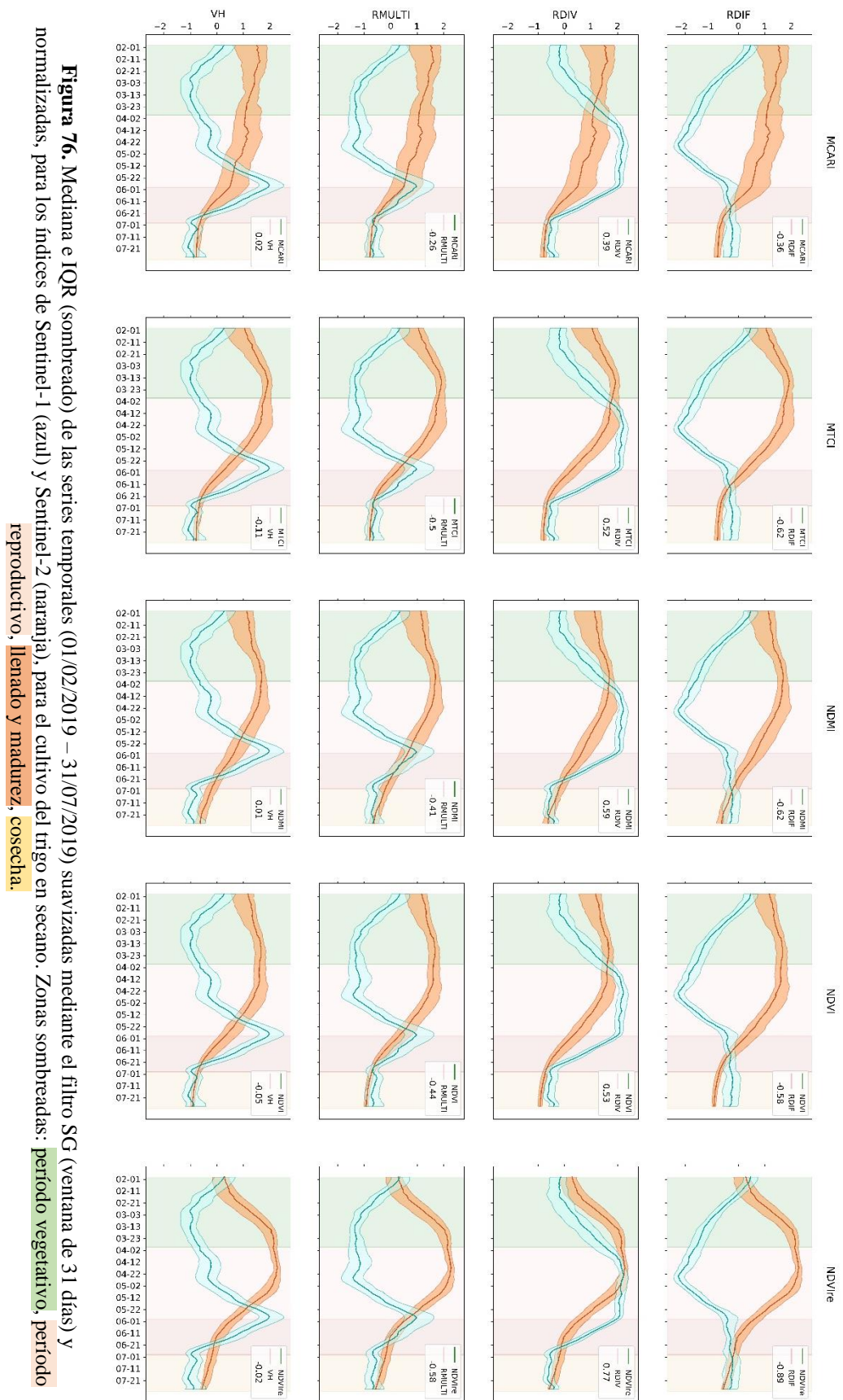


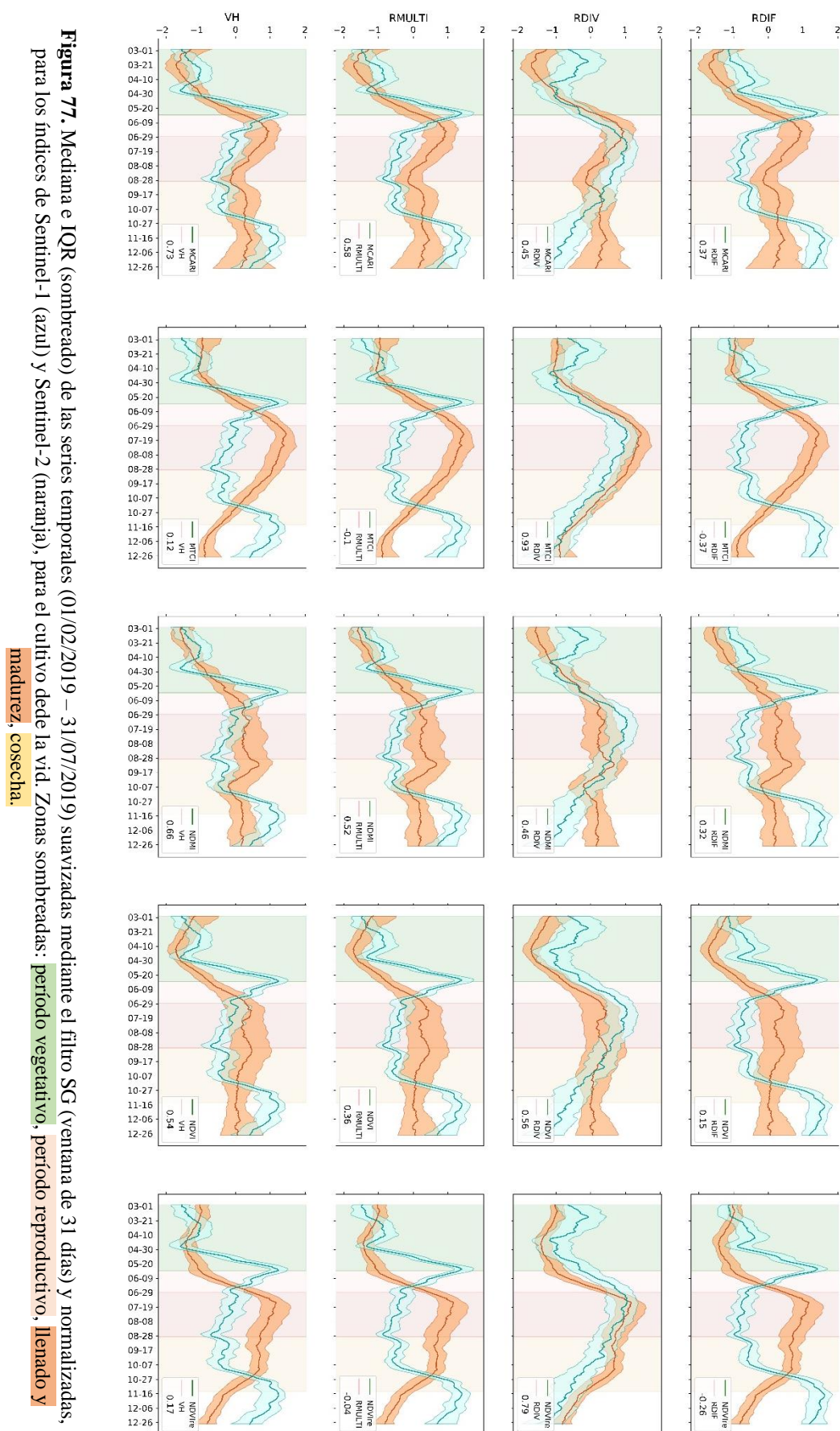




**Figura 75.** Mediana e IQR (sombreado) de las series temporales (01/02/2019 – 31/07/2019) suavizadas mediante el filtro SG (ventana de 31 días) y normalizadas, para los índices de Sentinel-1 (azul) y Sentinel-2 (naranja), para el cultivo del trigo en regadío. Zonas sombreadas: período vegetativo, período reproductivo, llenado y madurez, cosecha.

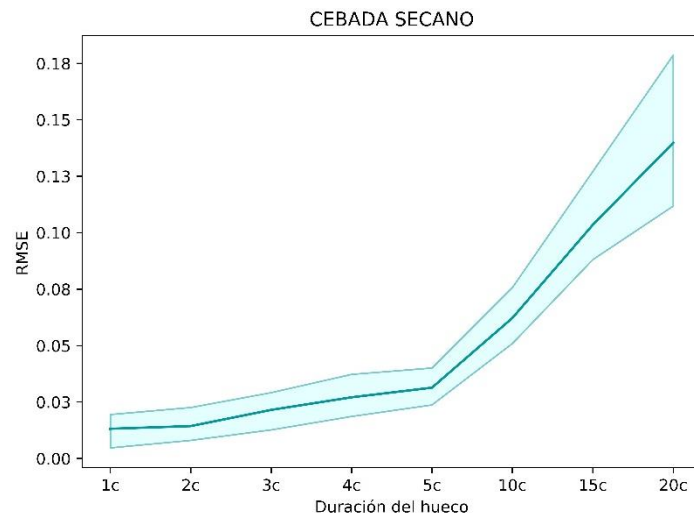




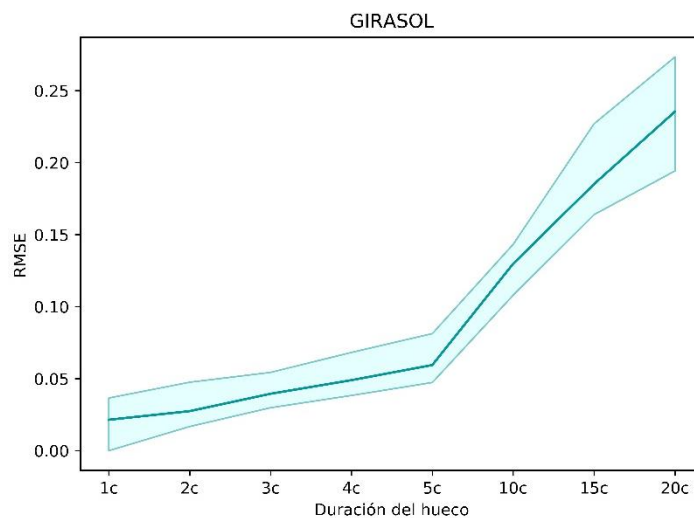


## ANEXO IV. RMSE DE EVALUACIÓN DE IMPACTO DE HUECOS

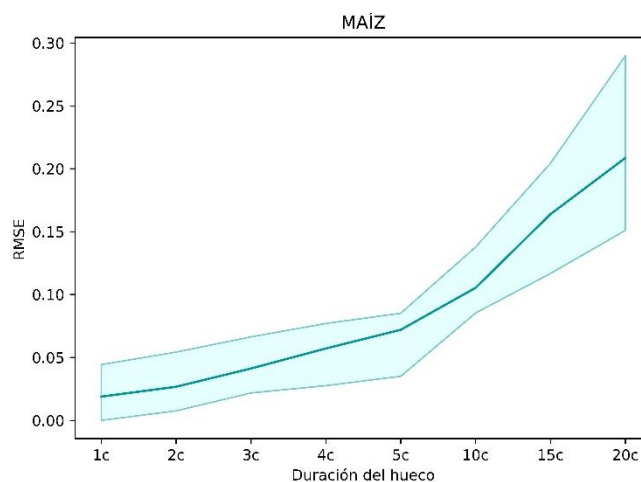
### EN FUNCIÓN DE LA DURACIÓN DEL HUECO



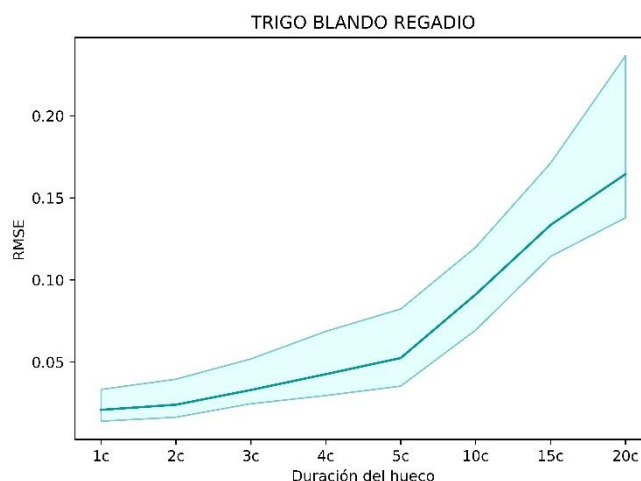
**Figura 78.** Representación de la mediana e IQR (sombreado) del RMSE de las parcelas del CT del cultivo de la cebada en secano en función de la duración de los huecos en las fechas de captura lo largo de las ventanas temporales definidas, para el índice NDVI.



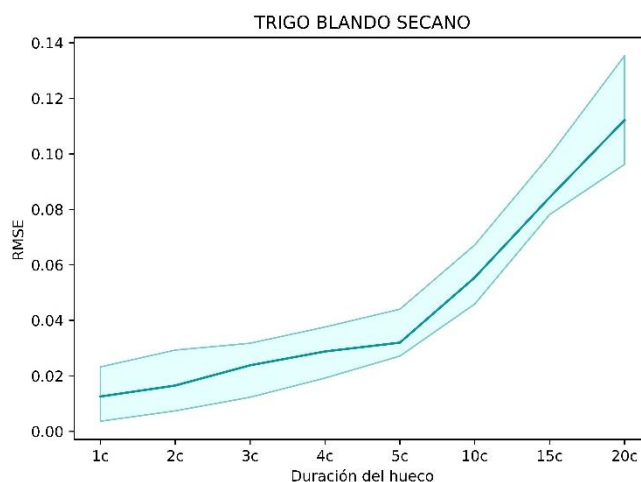
**Figura 79.** Representación de la mediana e IQR (sombreado) del RMSE de las parcelas del CT del cultivo del girasol en función de la duración de los huecos en las fechas de captura lo largo de las ventanas temporales definidas, para el índice NDVI.



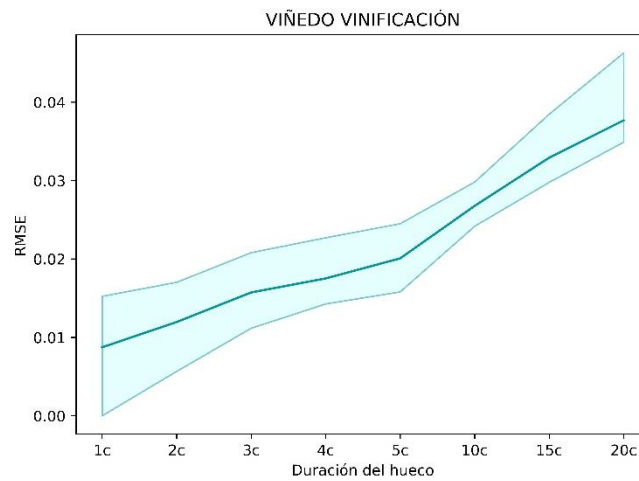
**Figura 80.** Representación de la mediana e IQR (sombreado) del RMSE de las parcelas del CT del cultivo del maíz en función de la duración de los huecos en las fechas de captura a lo largo de las ventanas temporales definidas, para el índice NDVI



**Figura 81.** Representación de la mediana e IQR (sombreado) del RMSE de las parcelas del CT del cultivo del trigo en regadío en función de la duración de los huecos en las fechas de captura a lo largo de las ventanas temporales definidas, para el índice NDVI



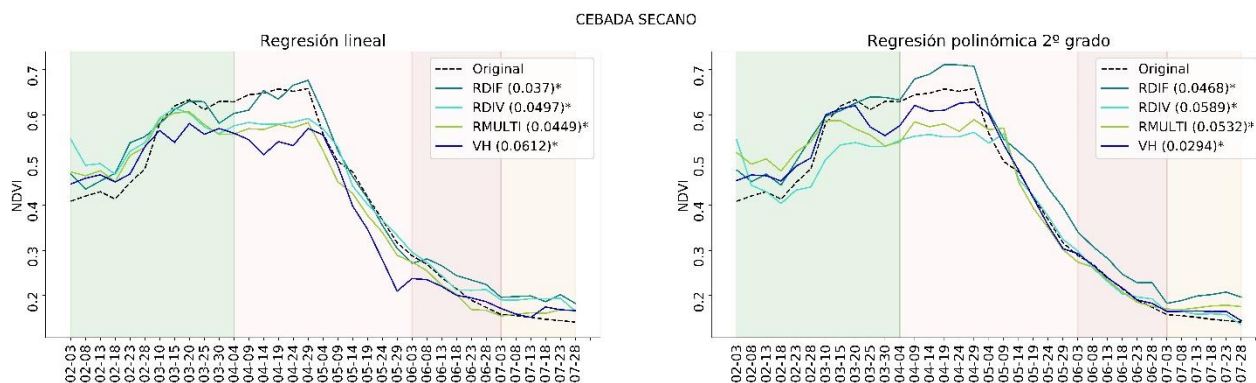
**Figura 82.** Representación de la mediana e IQR (sombreado) del RMSE de las parcelas del CT del cultivo del trigo en secano en función de la duración de los huecos en las fechas de captura a lo largo de las ventanas temporales definidas, para el índice NDVI.



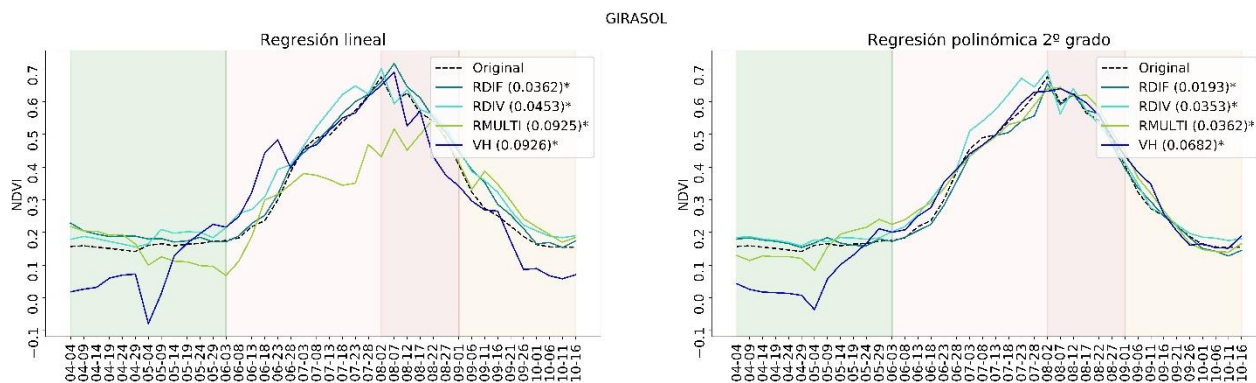
**Figura 83.** Representación de la mediana e IQR (sombreado) del RMSE de las parcelas del CT del cultivo de la vid en función de la duración de los huecos en las fechas de captura a lo largo de las ventanas temporales definidas, para el índice NDVI.



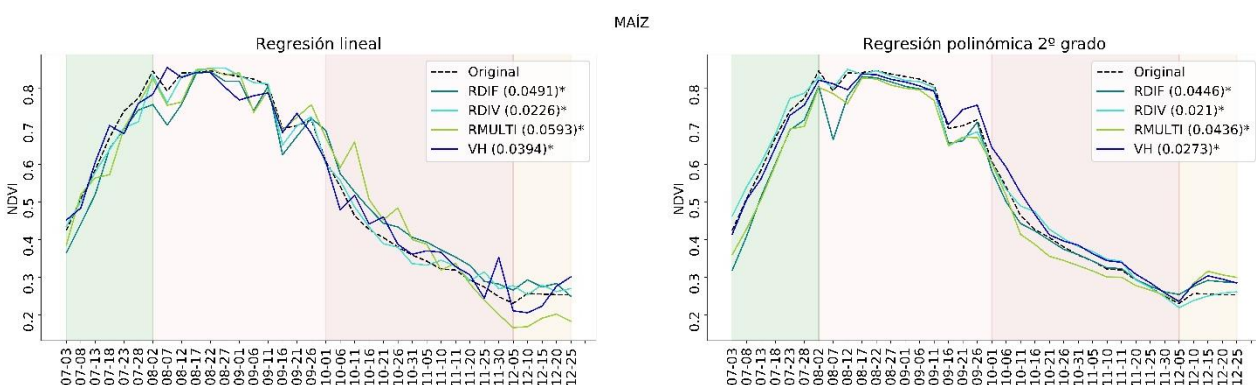
## ANEXO V. SERIES DE NDVI DE REFERENCIA PARA EL RELLENO DE HUECOS A PARTIR DE DATOS DE SENTINEL-1



**Figura 84.** Serie temporal NDVI de referencia obtenida a partir de datos de índices Sentinel-1 mediante regresión lineal (izquierda) y regresión polinómica de segundo grado (derecha) para el cultivo de la cebada en seco. Zonas sombreadas: **germinación y período vegetativo**, período reproductivo, **llenado y madurez**, cosecha.

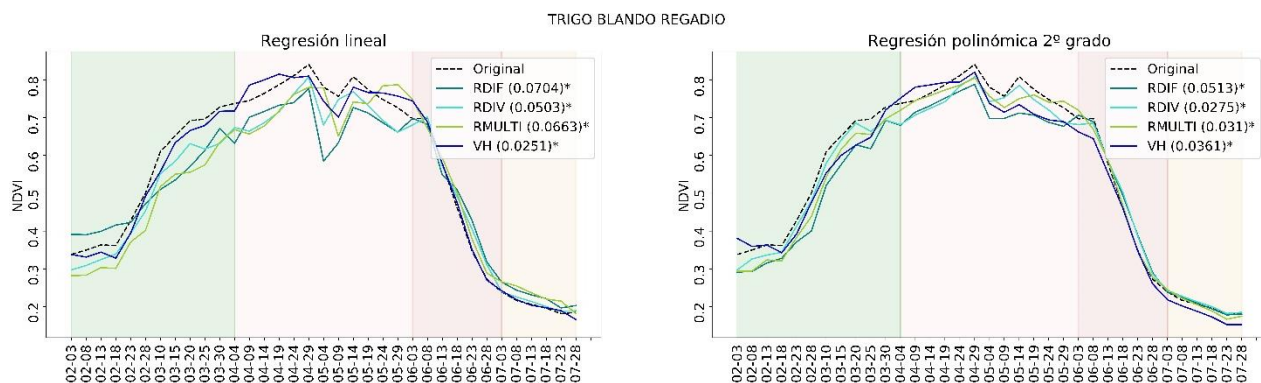


**Figura 85.** Serie temporal NDVI de referencia obtenida a partir de datos de índices Sentinel-1 mediante regresión lineal (izquierda) y regresión polinómica de segundo grado (derecha) para el cultivo del girasol. Zonas sombreadas: **germinación y período vegetativo**, período reproductivo, **llenado y madurez**, cosecha.

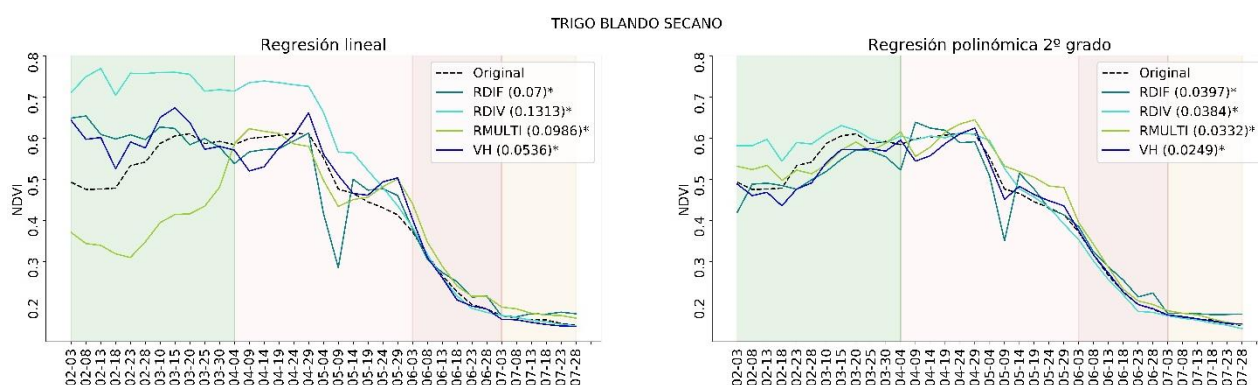


**Figura 86.** Serie temporal NDVI de referencia obtenida a partir de datos de índices Sentinel-1 mediante regresión lineal (izquierda) y regresión polinómica de segundo grado (derecha) para el cultivo del maíz. Zonas sombreadas: **germinación y período vegetativo**, período reproductivo, **llenado y madurez**, cosecha.

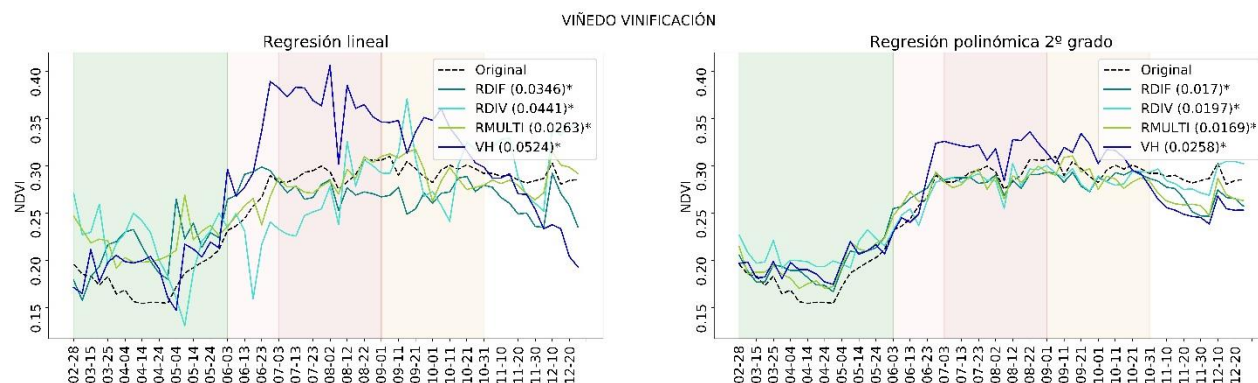




**Figura 87.** Serie temporal NDVI de referencia obtenida a partir de datos de índices Sentinel-1 mediante regresión lineal (izquierda) y regresión polinómica de segundo grado (derecha) para el cultivo del trigo en regadío. Zonas sombreadas: **germinación y período vegetativo**, **período reproductivo**, **llenado y madurez**, **cosecha**.



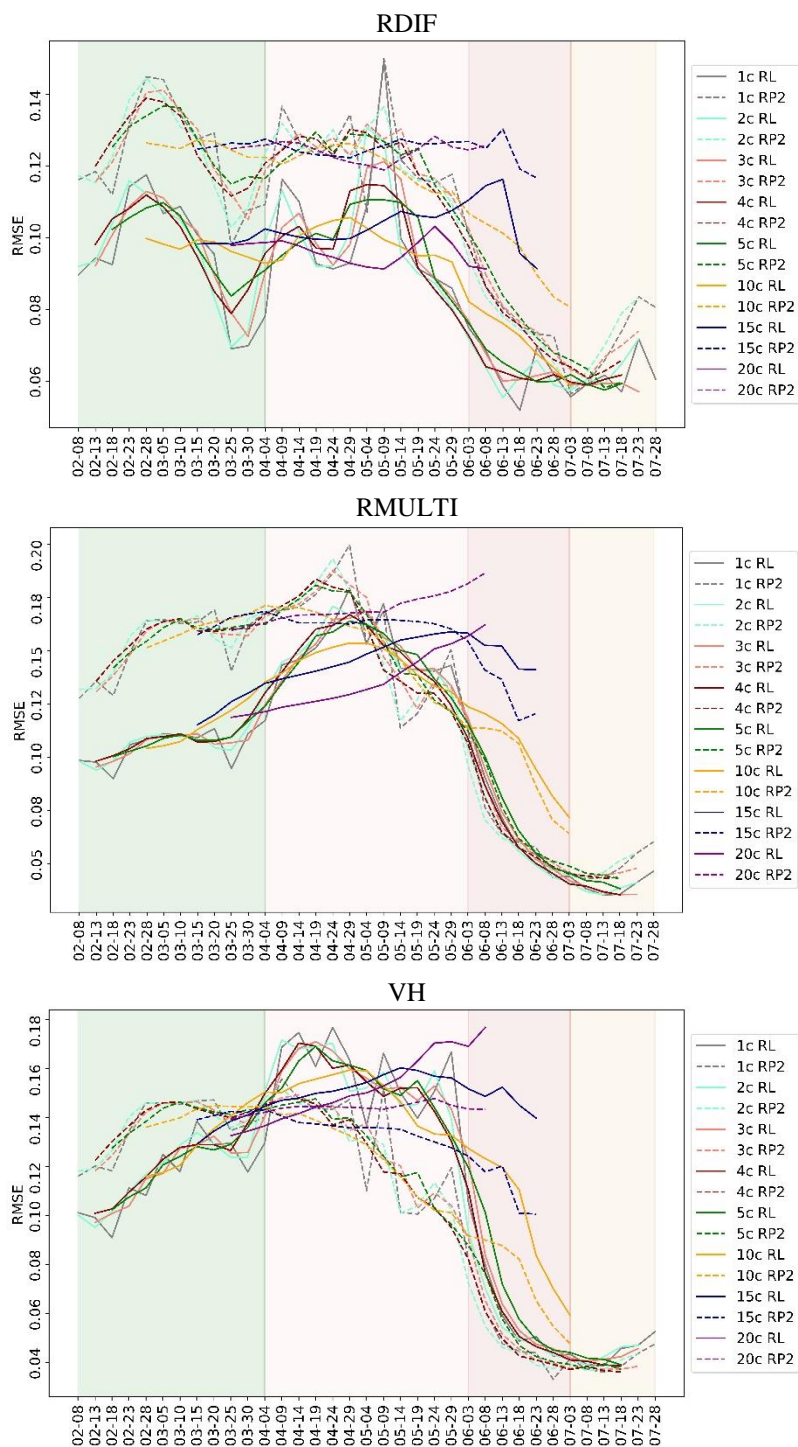
**Figura 88.** Serie temporal NDVI de referencia obtenida a partir de datos de índices Sentinel-1 mediante regresión lineal (izquierda) y regresión polinómica de segundo grado (derecha) para el cultivo del trigo en secano. Zonas sombreadas: **germinación y período vegetativo**, **período reproductivo**, **llenado y madurez**, **cosecha**.



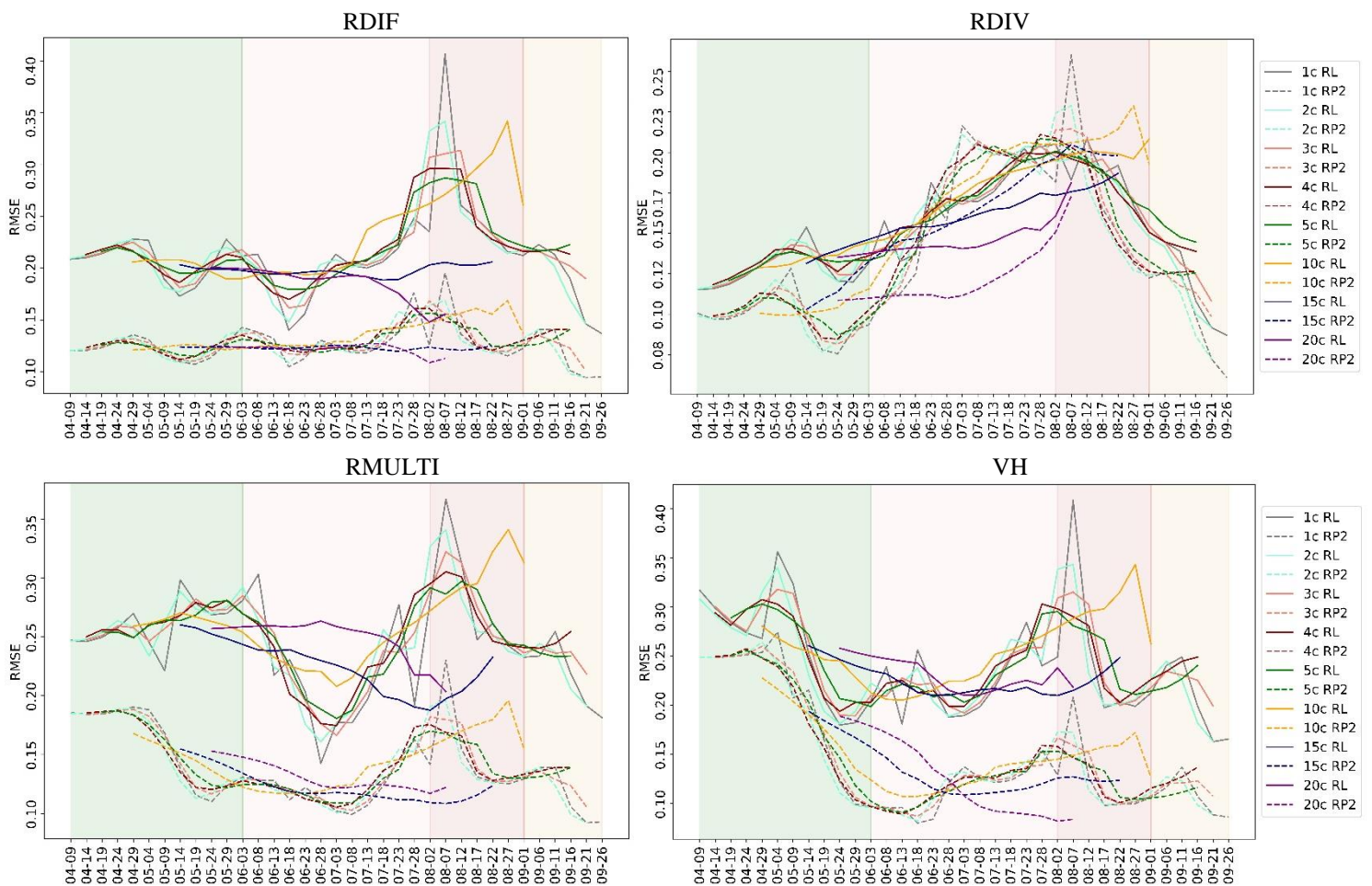
**Figura 89.** Serie temporal NDVI de referencia obtenida a partir de datos de índices Sentinel-1 mediante regresión lineal (izquierda) y regresión polinómica de segundo grado (derecha) para el cultivo de la vid. Zonas sombreadas: **germinación y período vegetativo**, **período reproductivo**, **llenado y madurez**, **cosecha**.

## ANEXO VI. RMSE DE LAS TÉCNICAS DE RELLENO DE HUECOS

### RMSE DE LAS TÉCNICAS DE RELLENO OBTENIDAS A PARTIR DE DATOS DE SENTINEL-1

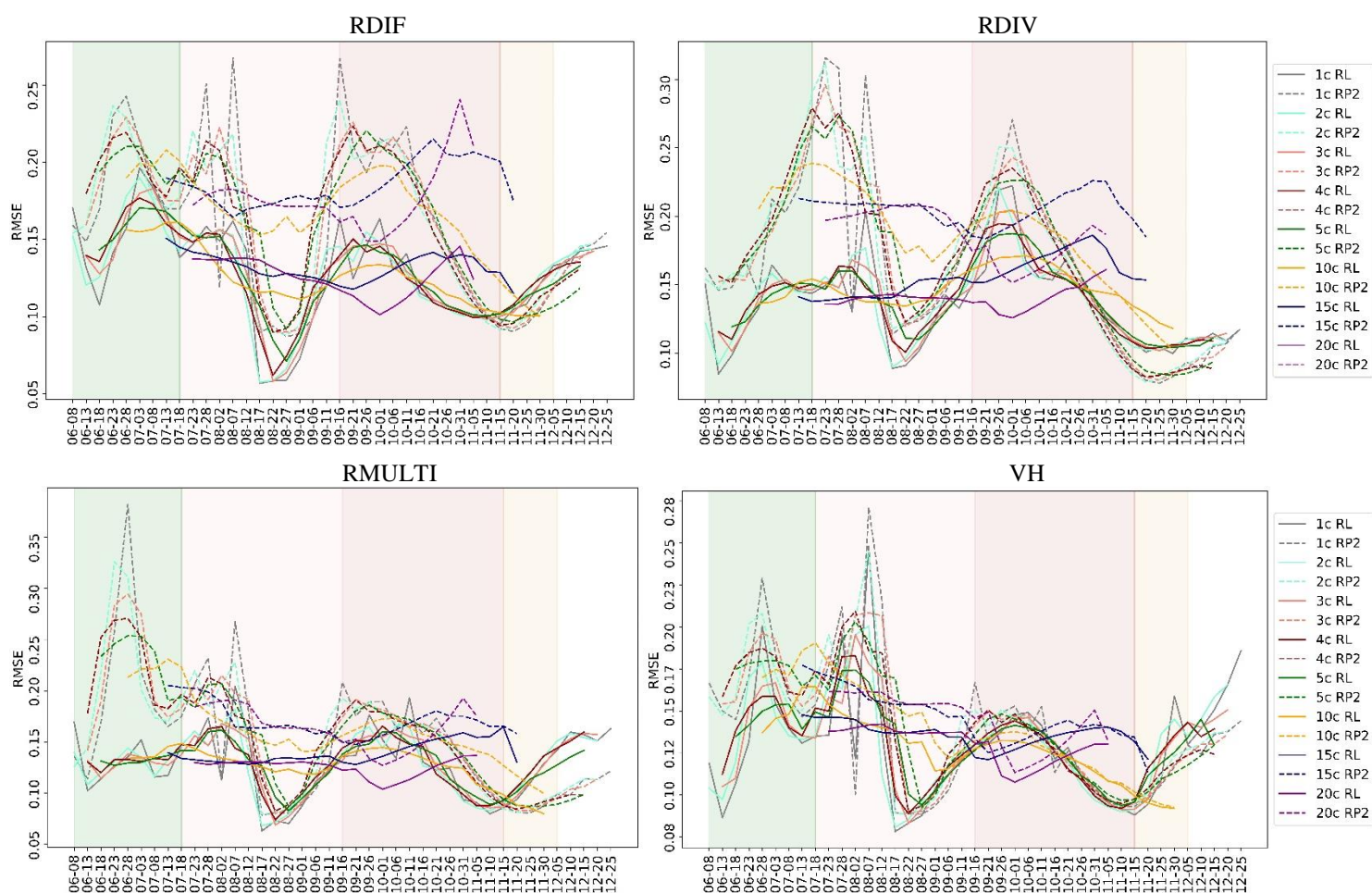


**Figura 90.** RMSE medio entre la serie de NDVI original ( $ST_i$ ) y las series reconstruidas ( $STAR_{S1}$ ) a partir de índices radar mediante regresiones lineales (RL, línea continua) y polinómicas (RP2, línea discontinua) para la cebada en seco, en función de la duración de los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

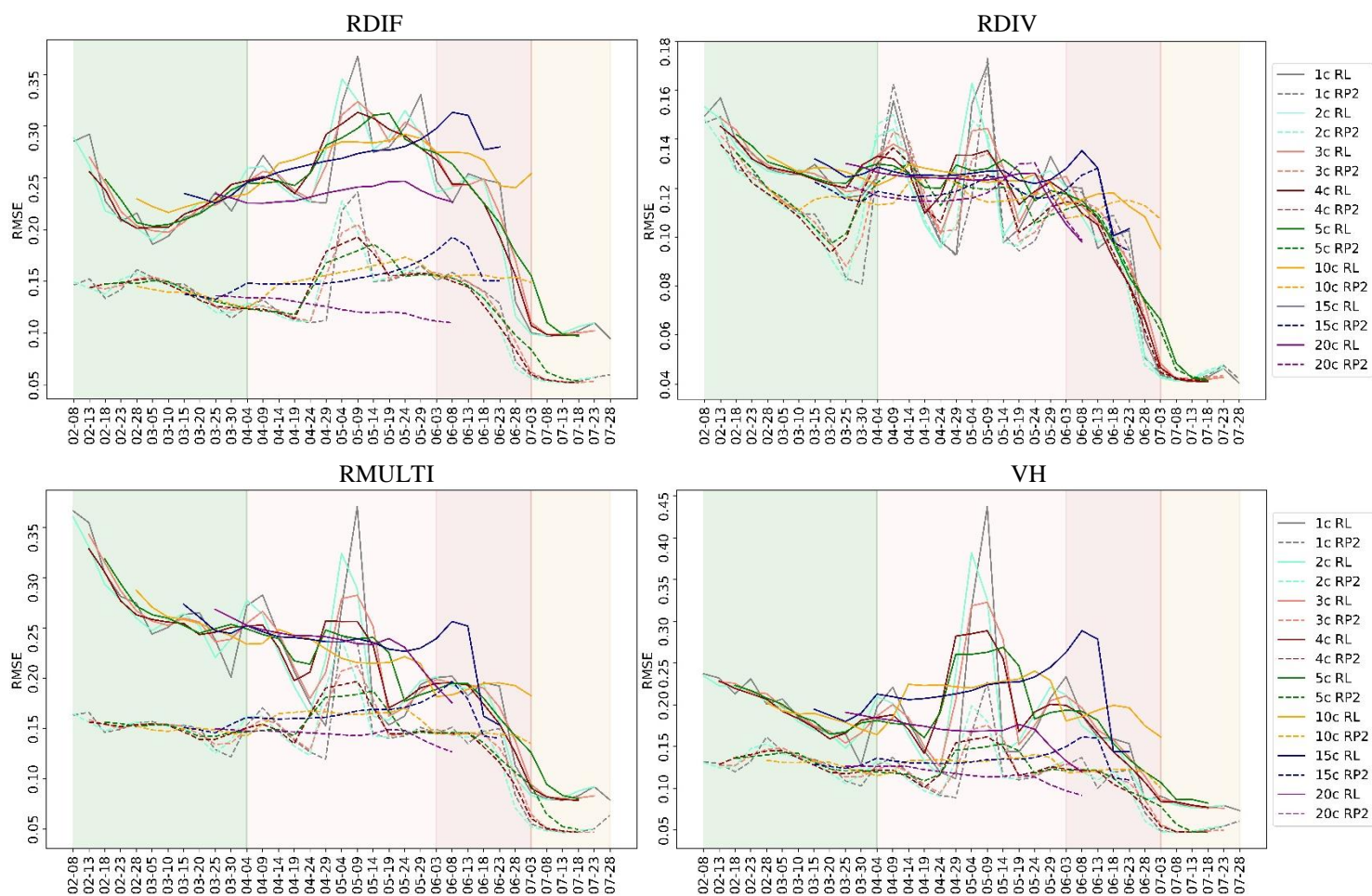


**Figura 91.** RMSE medio entre la serie de NDVI original ( $ST_i$ ) y las series reconstruidas ( $STAR_{S_i}$ ) a partir de índices radar mediante regresiones lineales (RL, línea continua) y polinómicas (RP2, línea discontinua) para el girasol, en función de la duración de los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

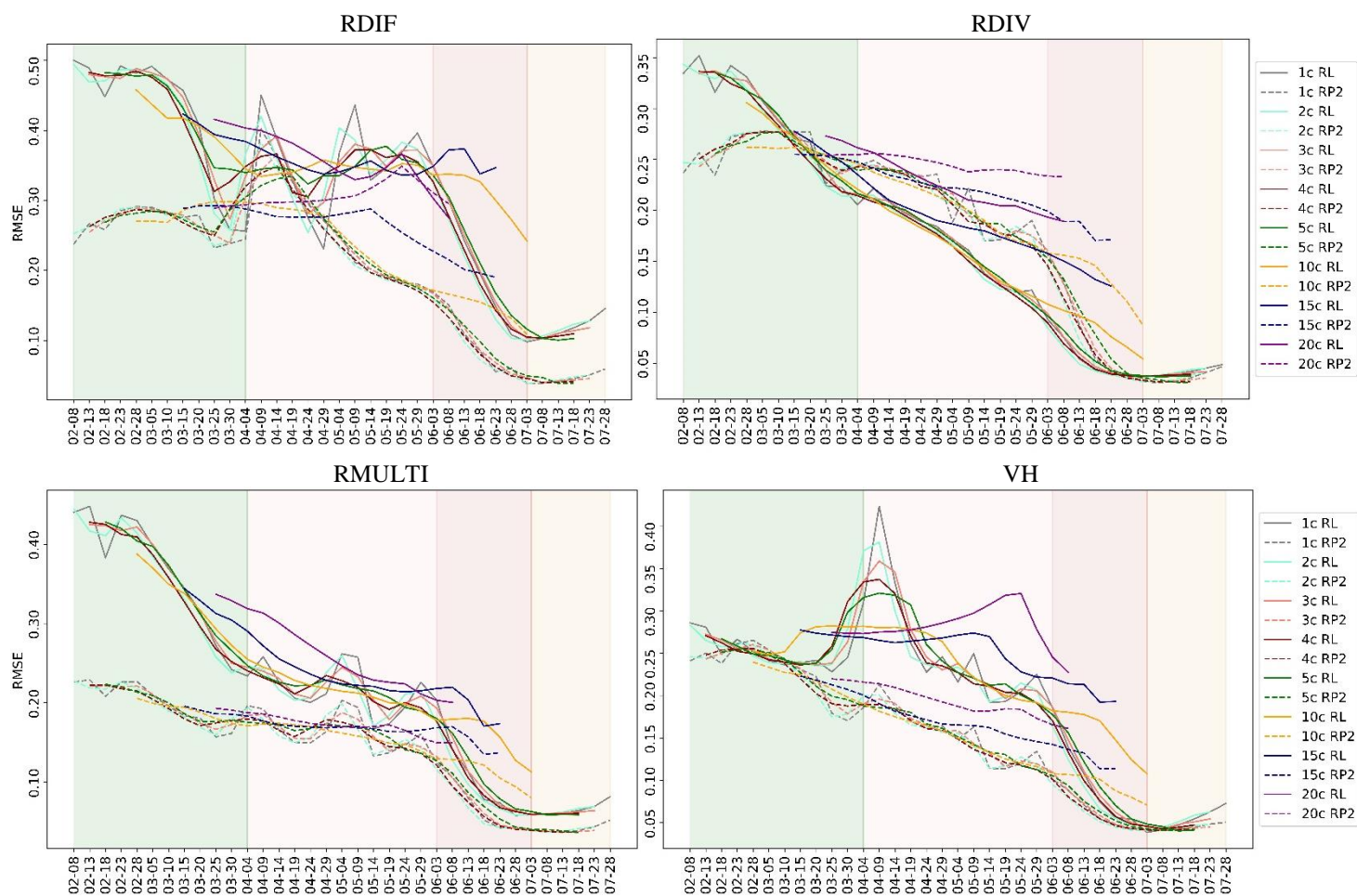




**Figura 92.** RMSE medio entre la serie de NDVI original (STi) y las series reconstruidas (STAR<sub>Si</sub>) a partir de índices radar mediante regresiones lineales (RL, línea continua) y polinómicas (RP2, línea discontinua) para el maíz, en función de la duración de los huecos. Zonas sombreadas: **germinación y período vegetativo**, **período reproductivo**, **llenado y madurez**, **cosecha**.

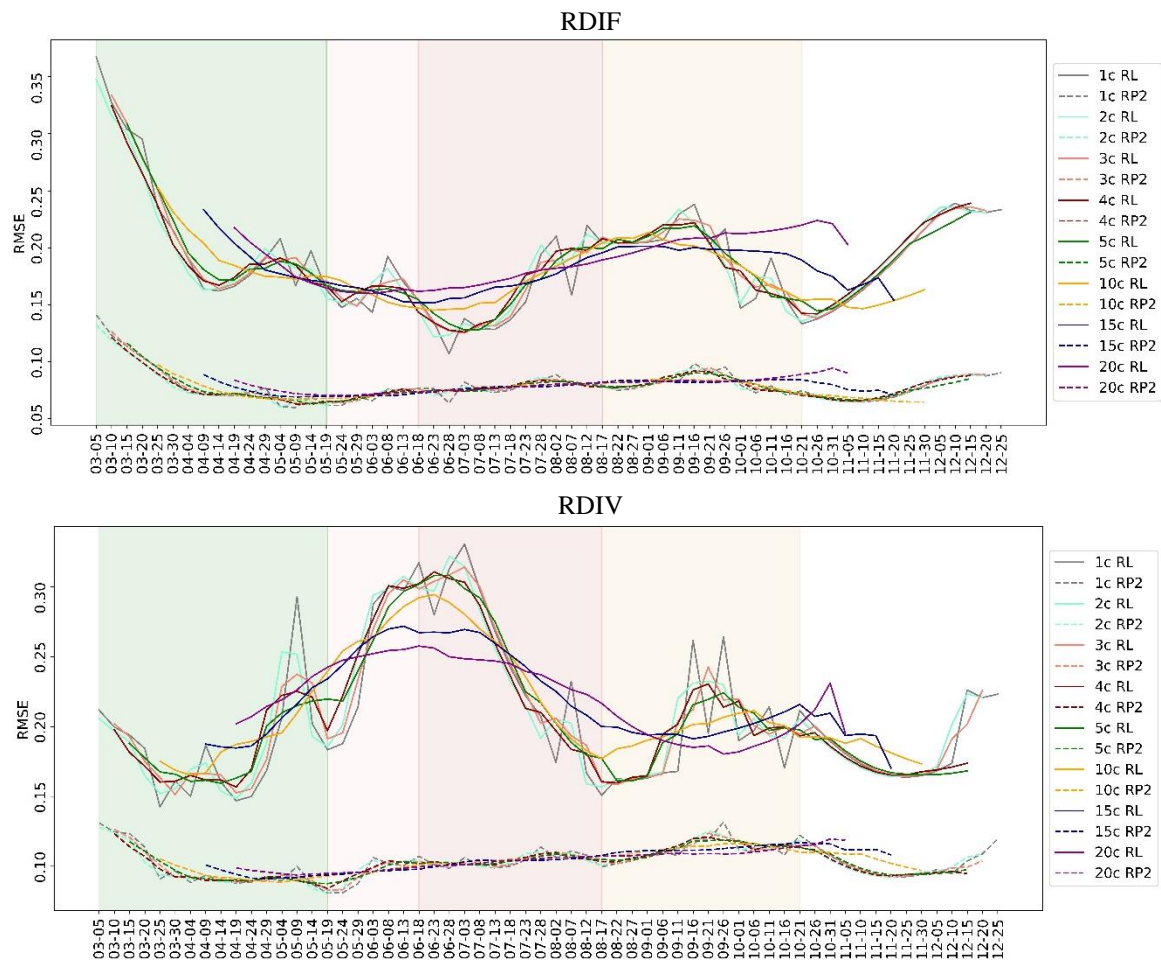


**Figura 93.** RMSE medio entre la serie de NDVI original (STi) y las series reconstruidas ( $STAR_{S1}$ ) a partir de índices radar mediante regresiones lineales (RL, línea continua) y polinómicas (RP2, línea discontinua) para el trigo en regadío, en función de la duración de los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

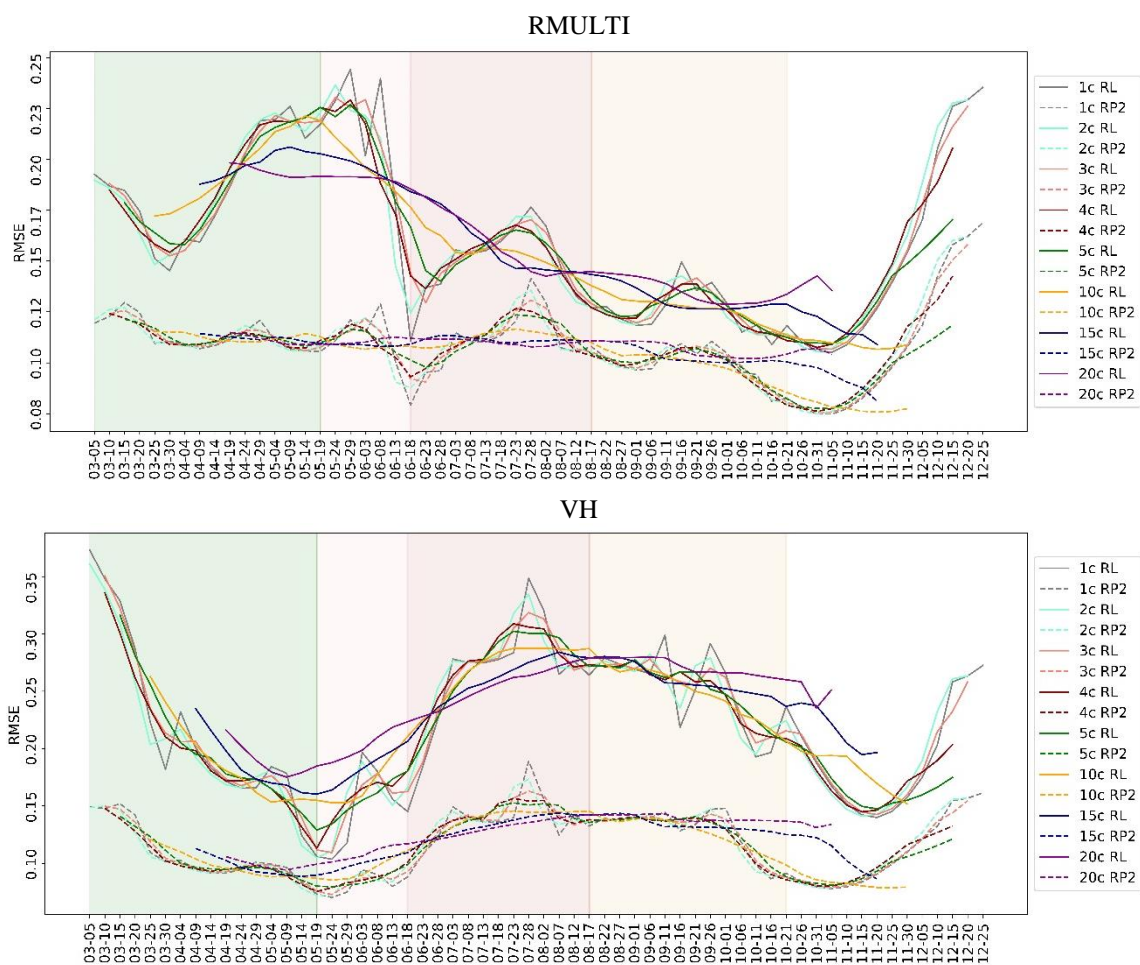


**Figura 94.** RMSE medio entre la serie de NDVI original (STi) y las series reconstruidas (STAR<sub>Si</sub>) a partir de índices radar mediante regresiones lineales (RL, línea continua) y polinómicas (RP2, línea discontinua) para el trigo en secano, en función de la duración de los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.





**Figura 95.** RMSE medio entre la serie de NDVI original ( $ST_i$ ) y las series reconstruidas ( $STAR_{S1}$ ) a partir de índices RDIF y RDIV mediante regresiones lineales (RL, línea continua) y polinómicas (RP2, línea discontinua) para la vid en función de la duración de los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.



**Figura 96.** RMSE medio entre la serie de NDVI original (STi) y las series reconstruidas (STAR<sub>Si</sub>) a partir de índices RMULTI y VH mediante regresiones lineales (RL, línea continua) y polinómicas (RP2, línea discontinua) para la vid en función de la duración de los huecos. Zonas sombreadas: germinación y período vegetativo, período reproductivo, llenado y madurez, cosecha.

## ANEXO VII. SCRIPTS DE PYTHON

### Script 1. Recorte de la zona de estudio

```
from osgeo import gdal
import numpy as np
import os
import errno
import time

#Definir la ruta de la zona de estudio:
zona =
r'D:\Escritorio\UPNA\BecaColaboracion\Datos\DatosModificados\Zona_rectangulo.shp'
#Definir nuevo directorio donde irán las carpetas recortadas:
nuevudir =
r'D:\Escritorio\UPNA\BecaColaboracion\Datos\DatosModificados\S2_r\2018\Recortes'
#Definir la ruta donde se encuentran las carpetas con los archivos a recortar:
##Formato: año/mes/día
diraño = r'D:\Escritorio\UPNA\BecaColaboracion\Datos\DatosOriginales\S2\2018'
carpetasaño = os.listdir(diraño)
for mes in carpetasaño:
    dirmes = diraño + '\\' + mes
    carpetasmes = os.listdir(dirmes)
    for carpeta in carpetasmes:
        #Definir el directorio de la carpeta de las imagenes a recortar:
        dircarpetai = dirmes + '\\' + carpeta
        os.chdir(dircarpetai)
        #Crear nueva carpeta donde irán los recortes:
        nuevacarpeta = nuevudir + '\\' + mes + carpeta + '_r'
        try:
            os.mkdir(nuevacarpeta)
        except OSError as e:
            if e.errno != errno.EEXIST:
                raise
    print('Recortando rasters de carpeta',carpeta)
    print (time.strftime("%d/%m/%y"),',', time.strftime("%H:%M:%S"))
    #Leer los rasters que hay en la carpeta
    for rast in os.listdir(dircarpetai):
        if rast[-4:] == '.jp2':

            #Abrir el raster con el que vamos a trabajar
            rt = gdal.Open(rast)
            #Definir el nombre y ruta del nuevo archivo:
            rt_r = nuevacarpeta + '\\' + rast[:-4] + '_r.jp2'
            #Cortamos con el área de trabajo:
            rt_z = gdal.Warp(rt_r, rt, cutlineDSName = zona,
                             cropToCutline = True, dstNodata = np.nan)
            rt = rt_r = rt_z = None
    print('Rasters de carpeta',carpeta,' recortados.')
    print (time.strftime("%d/%m/%y"),',', time.strftime("%H:%M:%S"))
```

### Script 2. Cálculo de índices Sentinel-1 y cálculo de estadísticas zonales.

```
"""
CÁLCULO ÍNDICES SENTINEL-1: RVI VV, VDDPI, IDPDD, Ratios (VH*VV, VV-VH,
VV+VH) .
"""
import os
import errno
import numpy as np
import gdal
import rasterio
import rasterstats
import geopandas as gdp
import pandas as pd
#import matplotlib.pyplot as plt
import time

dir0 = r'D:\Escritorio\UPNA\BecaColaboracion\Datos\DatosModificados\S1_r'
cont0 = os.listdir(dir0)
cont1 = []
for c in cont0:
    if '_' in c:
        cont1.append(c)
for carp in cont1:
    print('PROCESANDO PASADA',carp)
    #Leer las carpetas de la pasada correspondiente:
    dirpasada = dir0 + '\\' + carp
    os.chdir(dirpasada)
    contenidopasada0 = os.listdir(dirpasada)
    contenidopasada = []
    for cont in contenidopasada0:
        if '.xlsx' not in cont:
            contenidopasada.append(cont)

    ##Leer carpetas del año:
    for carpetas in contenidopasada:
        #Para cada carpeta a analizar:
        directorio = dirpasada + '\\' + carpetas
        os.chdir(directorio)
        contenido = os.listdir()
        print('Procesando carpeta',carpetas)
        print (time.strftime("%d/%m/%y"),',', time.strftime("%H:%M:%S"))
        #Encontrar las rutas donde está cada banda de interés
        for rast in contenido:
            #if 'B1' in rast and 'aux' not in rast:
            #    ruta_B1 = directorio + '\\' + rast
            if 'B2' in rast and 'aux' not in rast: ##Polarización cruzada:VH
                ruta_B2 = directorio + '\\' + rast
            if 'B3' in rast and 'aux' not in rast: ##Polarización directa: VV
                ruta_B3 = directorio + '\\' + rast
            #if 'B4' in rast and 'aux' not in rast:
            #    ruta_B4 = directorio + '\\' + rast

        #Abrir las bandas de interés:
```

```
#B1 = gdal.Open(ruta_B1)
B2 = gdal.Open(ruta_B2)
B3 = gdal.Open(ruta_B3)
#B4 = gdal.Open(ruta_B4)

#Leer las bandas como un array de numpy codificado como float de 32
bits:

#Datos_B1 = B1.GetRasterBand(1).ReadAsArray().astype(np.float32)
Datos_B2 = B2.GetRasterBand(1).ReadAsArray().astype(np.float32)
Datos_B3 = B3.GetRasterBand(1).ReadAsArray().astype(np.float32)
#Datos_B4 = B4.GetRasterBand(1).ReadAsArray().astype(np.float32)

#Convertir de dB a UL:
#B2
for i in range(Datos_B2.shape[0]-1):
    for j in range(Datos_B2.shape[1]-1):
        Datos_B2[i][j] = 10** (Datos_B2[i][j]/10)
#B3
for i in range(Datos_B3.shape[0]-1):
    for j in range(Datos_B3.shape[1]-1):
        Datos_B3[i][j] = 10** (Datos_B3[i][j]/10)

#Filtrar valores de VH < 0.32:
for i in range(Datos_B2.shape[0]):
    for j in range(Datos_B2.shape[1]):
        valor = Datos_B2[i][j]
        if valor > 0.32:
            Datos_B2[i][j] = 'nan'
        else:
            Datos_B2[i][j] = Datos_B2[i][j]

#Filtrar valores de VV < 1:
for i in range(Datos_B3.shape[0]):
    for j in range(Datos_B3.shape[1]):
        valor = Datos_B3[i][j]
        if valor > 1:
            Datos_B3[i][j] = 'nan'
        else:
            Datos_B3[i][j] = Datos_B3[i][j]

B3max = np.nanmax(Datos_B3)

#Extraer parámetros de interés para guardar los raster generados:
geotransformacion = B2.GetGeoTransform()
Xorigen, anchura_pixel, vacio, Yfinal, vacio2, altura_pixel =
geotransformacion
columnas = B2.RasterXSize
filas = B2.RasterYSize
proyeccion = B2.GetProjection()
Xfinal = Xorigen + (anchura_pixel * columnas)
Yorigen = Yfinal - (altura_pixel * filas)

#Crear una carpeta donde se almacenarán los productos intermedios:
```

```

carpetaintermedia = directorio + '\\ArchivosIntermedios'
try:
    os.mkdir(carpetaintermedia)
except OSError as e:
    if e.errno != errno.EEXIST:
        raise

#Crear una función para guardar los rasters creados con los índices:
def guardarRaster(dataset, datasetPath, columnas=columnas,
filas=filas, proyeccion=proyeccion):
    rasterSet = gdal.GetDriverByName('GTIFF').Create(datasetPath,
columnas, filas, 1, gdal.GDT_Float32)
    rasterSet.SetProjection(proyeccion)
    rasterSet.SetGeoTransform(geotransformacion)
    rasterSet.GetRasterBand(1).WriteArray(dataset)
    rasterSet.GetRasterBand(1).SetNoDataValue(0)
    rasterSet = None

#Crear función para calcular la media de los índices por parcela
zona =
gdp.read_file(r'D:\Escritorio\UPNA\BecaColaboracion\Datos\DatosModificados\zo
na_de_interés\Zona_B5m.shp')
def mediaParcela(img_array, affine, dir_xlsx, zona = zona):
    filas = len(zona)

    #Cálculo de los estadísticos zonales para cada raster
    med = rasterstats.zonal_stats(zona, img_array,
                                affine = affine,
                                stats = ['mean'],
                                geojason_out = True)

    #Crear DF de Pandas vacío sobre el que se reescribirán los datos:
    nombrecol = 'media'
    datosi = pd.DataFrame('', columns = [nombrecol], index =
np.arange(0,filas))
    datos = pd.DataFrame(index = np.arange(0,filas))
    #Insertar los datos de interés al DF de pandas
    for k in range(len(datos)):
        datosi.loc[i]['fid'] = zona.fid[i]
        datosi.loc[k][nombrecol] = med[k]['mean']

    #Concatenar los datos exsistentes con los nuevos:
    df = pd.concat([datos, datosi], axis=1)
    datos = df.copy()
    #Cálculo de los pixeles que hay por parcela:
    npix = rasterstats.zonal_stats(zona, img_array,
                                affine = affine,
                                stats = ['count'],
                                geojason_out = True)

    datos_np = pd.DataFrame('', columns = ['nº pixels'], index =
np.arange(0,filas))
    for j in range(len(datos)):
        datos_np.loc[j][['nº pixels']] = npix[j]['count']
    #Insertar una columna con el 'fid' identificativo de cada
parcela:

```



```

lista_fid = []
lista_desc = []
for z in range(len(datos)):
    lista_fid.append(zona.fid[z])
    lista_desc.append(zona.PRD_DESC[z])
datos.insert(loc = 0, column = 'fid', value = lista_fid)
datos.insert(loc = 1, column = 'cultivo', value = lista_desc)
datos.insert(loc = 2, column = 'nº pixels', value = npix)

#Generar un excel con el DF que se acaba de crear, guardandolo la
carpeta donde se encuentran los rasters:
datos.to_excel(dir_xlsx, header = True, index = False)

#B2:
B2_jp2 = rasterio.open(ruta_B2)
imgB20 = B2_jp2.read(1)
shape = imgB20.shape
imgB2 = np.empty(shape)
for i in range(imgB20.shape[0]-1):
    for j in range(imgB20.shape[1]-1):
        imgB2[i,j] = 10**(imgB20[i,j]/10)
affineB2 = B2_jp2.transform
nombre = str(ruta_B2)[96:114]
ruta_mediaB2 = carpetaintermedia + '\\' + nombre + '_mediaB2.xlsx'
mediaParcela(imgB2, affineB2, ruta_mediaB2)
df_B2 = pd.read_excel(ruta_mediaB2)
df_B2media = df_B2['media']
B2_np = df_B2media.to_numpy()
print('Media B2 calculada')
#print (time.strftime("%d/%m/%y"),',', time.strftime("%H:%M:%S"))

#B3:
B3_jp2 = rasterio.open(ruta_B3)
imgB30 = B3_jp2.read(1)
shape = imgB30.shape
imgB3 = np.empty(shape)
for i in range(imgB30.shape[0]-1):
    for j in range(imgB30.shape[1]-1):
        imgB3[i,j] = 10**(imgB30[i,j]/10)
affineB3 = B3_jp2.transform
nombre = str(ruta_B3)[96:114]
ruta_mediaB3 = carpetaintermedia + '\\' + nombre + '_mediaB3.xlsx'
mediaParcela(imgB3, affineB3, ruta_mediaB3)
df_B3 = pd.read_excel(ruta_mediaB3)
df_B3media = df_B3['media']
B3_np = df_B3media.to_numpy()
B3_npmax = np.nanmax(B3_np)
print('Media B3 calculada')
#print (time.strftime("%d/%m/%y"),',', time.strftime("%H:%M:%S"))

#Definición de las funciones para las dos opciones de cálculo:
##OPCIÓN A) Calcular primero el índice y luego la media por parcela
(el índice se calcula previo a utilizar la función)
def opcionA(INDICE_A, nombretiff, nombrexlsx, nombrecol):

```

```

global df_INDICE_A
#Convertir a raster:
ruta_INDICE_A_raster = carpetaintermedia + '\\\' + nombre + \'_\' +
nombretiff
guardarRaster(INDICE_A, ruta_INDICE_A_raster)
#Calcular la media del índice por parcela:
INDICE_tif = rasterio.open(ruta_INDICE_A_raster)
ruta_mediaINDICE = carpetaintermedia + '\\\' + nombre + \'_\' +
nombrexlsx
mediaParcela(INDICE_tif, ruta_mediaINDICE)
df_INDICE_A = pd.read_excel(ruta_mediaINDICE)
df_INDICE_A.columns = ['fid', 'cultivo', 'nº pixels', nombrecol]

#OPCIÓN B) Calcular los índices a partir de las medias de parcela (el
índice se calcula previo a utilizar la función)
def opcionB(INDICE_B, nombrecol):
    global df_INDICE_B
    df_INDICE_B = pd.DataFrame(INDICE_B)
    listafid = df_B2['fid']
    df_INDICE_B.insert(loc = 0, column = 'fid', value = listafid)
    df_INDICE_B.columns = ['fid', nombrecol]

#UNIÓN OPCIÓN A Y B:
def unionAB(df_INDICE_A, df_INDICE_B, INDICE, nombrexlsx):
    global df_INDICE, df_corr
    df_INDICE = df_INDICE_A.merge(df_INDICE_B, left_on = 'fid',
right_on = 'fid')
    df_corr = df_INDICE.corr()
    ruta_INDICE = directorio + '\\\' + nombre + \'_\' + nombrexlsx
    writer = pd.ExcelWriter(ruta_INDICE, engine = 'xlsxwriter')
    df_INDICE.to_excel(writer, sheet_name = INDICE + \'_AB\')
    df_corr.to_excel(writer, sheet_name = 'Correlación')
    writer.save()
    writer.close()

##### CÁLCULO DE ÍNDICES #####

#### RVI ####
def RVI(Datos_B2, Datos_B3):
    return np.divide(4*Datos_B2, Datos_B3 + Datos_B2, where =
(Datos_B3 + Datos_B2) != 0)
    RVI_A = RVI(Datos_B2, Datos_B3)
    RVI_B = RVI(B2_np, B3_np)
    opcionA(RVI_A, 'RVI_A.tiff', 'RVI_A.xlsx', 'RVI_A')
    opcionB(RVI_B, 'RVI_B')
    unionAB(df_INDICE_A, df_INDICE_B, 'RVI', 'RVI.xlsx')
    print('Índice RVI calculado.')

#### VDDPI ####
def VDDPI(Datos_B2, Datos_B3):
    return np.divide(Datos_B2 + Datos_B3, Datos_B3, where =
(Datos_B3) != 0)
    VDDPI_A = VDDPI(Datos_B2, Datos_B3)
    VDDPI_B = VDDPI(B2_np, B3_np)

```

```
opcionA(VDDPI_A, 'VDDPI_A.tiff', 'VDDPI_A.xlsx', 'VDDPI_A')
opcionB(VDDPI_B, 'VDDPI_B')
unionAB(df_INDICE_A, df_INDICE_B, 'VDDPI', 'VDDPI.xlsx')
print('Índice VDDPI calculado.')

#### IDPDD ####
def IDPDD(Datos_B2, Datos_B3, B3max):
    return np.divide((B3max-Datos_B3) + Datos_B2, 2**0.5)
IDPDD_A = IDPDD(Datos_B2, Datos_B3, B3max)
IDPDD_B = IDPDD(B2_np, B3_np, B3_npmax)
opcionA(IDPDD_A, 'IDPDD_A.tiff', 'IDPDD_A.xlsx', 'IDPDD_A')
opcionB(IDPDD_B, 'IDPDD_B')
unionAB(df_INDICE_A, df_INDICE_B, 'IDPDD', 'IDPDD.xlsx')
print('Índice IDPDD calculado.')

#### RDIV ####
def RDIV(Datos_B2, Datos_B3):
    return np.divide(Datos_B2, Datos_B3, where = (Datos_B3) != 0)
RDIV_A = RDIV(Datos_B2, Datos_B3)
RDIV_B = RDIV(B2_np, B3_np)
opcionA(RDIV_A, 'RDIV_A.tiff', 'RDIV_A.xlsx', 'RDIV_A')
opcionB(RDIV_B, 'RDIV_B')
unionAB(df_INDICE_A, df_INDICE_B, 'RDIV', 'RDIV.xlsx')
print('Índice RDIV calculado.')

#### RMULTI ####
def RMULTI(Datos_B2, Datos_B3):
    return (Datos_B2 * Datos_B3)
RMULTI_A = RMULTI(Datos_B2, Datos_B3)
RMULTI_B = RMULTI(B2_np, B3_np)
opcionA(RMULTI_A, 'RMULTI_A.tiff', 'RMULTI_A.xlsx', 'RMULTI_A')
opcionB(RMULTI_B, 'RMULTI_B')
unionAB(df_INDICE_A, df_INDICE_B, 'RMULTI', 'RMULTI.xlsx')
print('Índice RMULTI calculado.')

#### RDIF ####
def RDIF(Datos_B2, Datos_B3):
    return (Datos_B3 - Datos_B2)
RDIF_A = RDIF(Datos_B2, Datos_B3)
RDIF_B = RDIF(B2_np, B3_np)
opcionA(RDIF_A, 'RDIF_A.tiff', 'RDIF_A.xlsx', 'RDIF_A')
opcionB(RDIF_B, 'RDIF_B')
unionAB(df_INDICE_A, df_INDICE_B, 'RDIF', 'RDIF.xlsx')
print('Índice RDIF calculado.')

#### RSUM ####
def RSUM(Datos_B2, Datos_B3):
    return (Datos_B2 + Datos_B3)
RSUM_A = RSUM(Datos_B2, Datos_B3)
RSUM_B = RSUM(B2_np, B3_np)
opcionA(RSUM_A, 'RSUM_A.tiff', 'RSUM_A.xlsx', 'RSUM_A')
opcionB(RSUM_B, 'RSUM_B')
unionAB(df_INDICE_A, df_INDICE_B, 'RSUM', 'RSUM.xlsx')
print('Índice RSUM calculado.')
```

```
#### VH ####
def VH(Datos_B2):
    return (Datos_B2)
VH_A = VH(Datos_B2)
VH_B = VH(B2_np)
opcionA(VH_A, 'VH_A.tiff', 'VH_A.xlsx', 'VH_A')
opcionB(VH_B, 'VH_B')
unionAB(df_INDICE_A, df_INDICE_B, 'VH', 'VH.xlsx')
print('Índice VH calculado.')

#### VV ####
def VV(Datos_B3):
    return (Datos_B3)
VV_A = VV(Datos_B3)
VV_B = VV(B3_np)
opcionA(VV_A, 'VV_A.tiff', 'VV_A.xlsx', 'VV_A')
opcionB(VV_B, 'VV_B')
unionAB(df_INDICE_A, df_INDICE_B, 'VV', 'VV.xlsx')
print('Índice VV calculado.')
```

### Script 3. Cálculo de índices de Sentinel-2 y estadísticas zonales.

```
# -*- coding: utf-8 -*-
"""
CÁLCULO ÍNDICES SENTINEL-2: NDVI, SAVI, RDVI, NDMI, SLAVI, CI-RE, MCARI, MTCI,
NDVIre y SeLI.
"""
import os
import errno
import numpy as np
import gdal
import rasterio
import rasterstats
import geopandas as gdp
import pandas as pd
import matplotlib.pyplot as plt
import time

#Leer las carpetas del año correspondiente:
diraño =
r'D:\Escritorio\UPNA\BecaColaboracion\Datos\DatosModificados\S2_r\2019'
os.chdir(diraño)
contenidoaño0 = os.listdir(diraño)
contenidoaño = []
for c in contenidoaño0:
    if '.xlsx' not in c:
        contenidoaño.append(c)

##Leer carpetas del año:
for carpetas in contenidoaño:
    #Para cada carpeta a analizar:
    directorio = diraña + '\\' + carpetas
    os.chdir(directorio)
    contenido = os.listdir()
    print('Procesando carpeta',carpetas)
    print (time.strftime("%d/%m/%y"),',', time.strftime("%H:%M:%S"))
    #Encontrar las rutas donde está cada banda de interés
    for rast in contenido:
        if 'B03' in rast and 'aux' not in rast:
            ruta_B3 = directorio + '\\' + rast
        if 'B04' in rast and 'aux' not in rast:
            ruta_B4 = directorio + '\\' + rast
        if 'B05' in rast and 'aux' not in rast:
            ruta_B5 = directorio + '\\' + rast
        if 'B06' in rast and 'aux' not in rast:
            ruta_B6 = directorio + '\\' + rast
        if 'B07' in rast and 'aux' not in rast:
            ruta_B7 = directorio + '\\' + rast
        if 'B08' in rast and 'aux' not in rast:
            ruta_B8 = directorio + '\\' + rast
        if 'B11' in rast and 'aux' not in rast:
            ruta_B11 = directorio + '\\' + rast
        if 'B12' in rast and 'aux' not in rast:
            ruta_B12 = directorio + '\\' + rast

#Abrir las bandas de interés:
```



```
B3 = gdal.Open(ruta_B3)
B4 = gdal.Open(ruta_B4)
B5 = gdal.Open(ruta_B5)
B6 = gdal.Open(ruta_B6)
B7 = gdal.Open(ruta_B7)
B8 = gdal.Open(ruta_B8)
B11 = gdal.Open(ruta_B11)
B12 = gdal.Open(ruta_B12)

#Leer las bandas como un array de numpy codificado como float de 32 bits:
Datos_B3 = B3.GetRasterBand(1).ReadAsArray().astype(np.int32)/10000
Datos_B4 = B4.GetRasterBand(1).ReadAsArray().astype(np.int32)/10000
Datos_B5 = B5.GetRasterBand(1).ReadAsArray().astype(np.int32)/10000
Datos_B6 = B6.GetRasterBand(1).ReadAsArray().astype(np.int32)/10000
Datos_B7 = B7.GetRasterBand(1).ReadAsArray().astype(np.int32)/10000
Datos_B8 = B8.GetRasterBand(1).ReadAsArray().astype(np.int32)/10000
Datos_B11 = B11.GetRasterBand(1).ReadAsArray().astype(np.int32)/10000
Datos_B12 = B12.GetRasterBand(1).ReadAsArray().astype(np.int32)/10000

#Extraer parámetros de interés para guardar los raster generados:
geotransformacion = B4.GetGeoTransform()
Xorigen, anchura_pixel, vacio, Yfinal, vacio2, altura_pixel =
geotransformacion
columnas = B4.RasterXSize
filas = B4.RasterYSize
proyeccion = B4.GetProjection()
Xfinal = Xorigen + (anchura_pixel * columnas)
Yorigen = Yfinal - (altura_pixel * filas)

#Crear una carpeta donde se almacenarán los productos intermedios:
carpetaintermedia = directorio + '\\ArchivosIntermedios'
try:
    os.mkdir(carpetaintermedia)
except OSError as e:
    if e.errno != errno.EEXIST:
        raise

#Crear una función para guardar los rasters creados con los índices:
def guardarRaster(dataset, datasetPath, columnas=columnas, filas=filas,
proyeccion=proyeccion):
    rasterSet = gdal.GetDriverByName('GTIFF').Create(datasetPath,
columnas, filas, 1, gdal.GDT_Float32)
    rasterSet.SetProjection(proyeccion)
    rasterSet.SetGeoTransform(geotransformacion)
    rasterSet.GetRasterBand(1).WriteArray(dataset)
    rasterSet.GetRasterBand(1).SetNoDataValue(0)
    rasterSet = None

#Crear función para calcular la media de los índices por parcela
zona =
gdp.read_file(r'D:\Escritorio\UPNA\BecaColaboracion\Datos\DatosModificados\zon
a_de_interés\Zona_B5m.shp')
def mediaParcela(img, dir_xlsx, zona = zona):
    filas = len(zona)
    #Pasar la información a un array de numpy
```

```
img_array = img.read(1)
affine = img.transform

#Cálculo de los estadísticos zonales para cada raster
med = rasterstats.zonal_stats(zona, img_array,
                              affine = affine,
                              stats = ['mean'],
                              geojason_out = True)

#Crear DF de Pandas vacío sobre el que se reescribirán los datos:
nombrecol = 'media'
datosi = pd.DataFrame('', columns = [nombrecol], index =
np.arange(0,filas))
datos = pd.DataFrame(index = np.arange(0,filas))
#Insertar los datos de interés al DF de pandas
for k in range(len(datos)):
    datosi.loc[i]['fid'] = zona.fid[i]
    datosi.loc[k][nombrecol] = med[k]['mean']

#Concatenar los datos exsistentes con los nuevos:
df = pd.concat([datos, datosi], axis=1)
datos = df.copy()
#Cálculo de los pixeles que hay por parcela:
npix = rasterstats.zonal_stats(zona, img_array,
                              affine = affine,
                              stats = ['count'],
                              geojason_out = True)

datos_np = pd.DataFrame('', columns = ['nº pixels'], index =
np.arange(0,filas))
for j in range(len(datos)):
    datos_np.loc[j][['nº pixels']] = npix[j]['count']
#Insertar una columna con el 'fid' identificativo de cada parcela:
lista_fid = []
lista_desc = []
for z in range(len(datos)):
    lista_fid.append(zona.fid[z])
    lista_desc.append(zona.PRDesc[z])
datos.insert(loc = 0, column = 'fid', value = lista_fid)
datos.insert(loc = 1, column = 'cultivo', value = lista_desc)
datos.insert(loc = 2, column = 'nº pixels', value = npix)

#Generar un excel con el DF que se acaba de crear, guardandolo la
carpeta donde se encuentran los rasters:
datos.to_excel(dir_xlsx, header = True, index = False)

#Obtener los valores medios de reflectancia por parcela para cada banda y
convertirlo a un array de numpy:
#B3:
B3_jp2 = rasterio.open(ruta_B3)
nombre = str(ruta_B3)[76:91]
ruta_mediaB3 = carpetaintermedia + '\\\\' + nombre + '_mediaB3.xlsx' #Se
guarda la media en unidades de reflectancia x 10000
#mediaParcela(B3_jp2, ruta_mediaB3)
df_B3 = pd.read_excel(ruta_mediaB3)
df_B3media = df_B3['media']/10000
B3_np = df_B3media.to_numpy()
```

```

print('Media B3 calculada')
print (time.strftime("%d/%m/%y"),',', time.strftime("%H:%M:%S"))
#B4:
B4_jp2 = rasterio.open(ruta_B4)
nombre = str(ruta_B4)[76:91]
ruta_mediaB4 = carpetaintermedia + '\\\\' + nombre + '_mediaB4.xlsx' #Se
guarda la media en unidades de reflectancia x 10000
#mediaParcela(B4_jp2, ruta_mediaB4)
df_B4 = pd.read_excel(ruta_mediaB4)
df_B4media = df_B4['media']/10000
B4_np = df_B4media.to_numpy()
print('Media B4 calculada')
print (time.strftime("%d/%m/%y"),',', time.strftime("%H:%M:%S"))
#B5:
B5_jp2 = rasterio.open(ruta_B5)
nombre = str(ruta_B5)[76:91]
ruta_mediaB5 = carpetaintermedia + '\\\\' + nombre + '_mediaB5.xlsx' #Se
guarda la media en unidades de reflectancia x 10000
#mediaParcela(B5_jp2, ruta_mediaB5)
df_B5 = pd.read_excel(ruta_mediaB5)
df_B5media = df_B5['media']/10000
B5_np = df_B5media.to_numpy()
print('Media B5 calculada')
print (time.strftime("%d/%m/%y"),',', time.strftime("%H:%M:%S"))
#B6:
B6_jp2 = rasterio.open(ruta_B6)
nombre = str(ruta_B6)[76:91]
ruta_mediaB6 = carpetaintermedia + '\\\\' + nombre + '_mediaB6.xlsx' #Se
guarda la media en unidades de reflectancia x 10000
#mediaParcela(B6_jp2, ruta_mediaB6)
df_B6 = pd.read_excel(ruta_mediaB6)
df_B6media = df_B6['media']/10000
B6_np = df_B6media.to_numpy()
print('Media B6 calculada')
print (time.strftime("%d/%m/%y"),',', time.strftime("%H:%M:%S"))
#B7:
B7_jp2 = rasterio.open(ruta_B7)
nombre = str(ruta_B7)[76:91]
ruta_mediaB7 = carpetaintermedia + '\\\\' + nombre + '_mediaB7.xlsx' #Se
guarda la media en unidades de reflectancia x 10000
#mediaParcela(B7_jp2, ruta_mediaB7)
df_B7 = pd.read_excel(ruta_mediaB7)
df_B7media = df_B7['media']/10000
B7_np = df_B7media.to_numpy()
print('Media B7 calculada')
print (time.strftime("%d/%m/%y"),',', time.strftime("%H:%M:%S"))
#B8:
B8_jp2 = rasterio.open(ruta_B8)
nombre = str(ruta_B8)[76:91]
ruta_mediaB8 = carpetaintermedia + '\\\\' + nombre + '_mediaB8.xlsx' #Se
guarda la media en unidades de reflectancia x 10000
#mediaParcela(B8_jp2, ruta_mediaB8)
df_B8 = pd.read_excel(ruta_mediaB8)
df_B8media = df_B8['media']/10000
B8_np = df_B8media.to_numpy()

```

```

print('Media B8 calculada')
print (time.strftime("%d/%m/%y"),',', time.strftime("%H:%M:%S"))
#B11:
#B11_jp2 = rasterio.open(ruta_B11)
#nombre = str(ruta_B11)[76:91]
#ruta_mediaB11 = carpetaintermedia + '\\\' + nombre + '_mediaB11.xlsx' #Se
guarda la media en unidades de reflectancia x 10000
#mediaParcela(B11_jp2, ruta_mediaB11)
#df_B11 = pd.read_excel(ruta_mediaB11)
#df_B11media = df_B11['media']/10000
#B11_np = df_B11media.to_numpy()
#print('Media B11 calculada')
#print (time.strftime("%d/%m/%y"),',', time.strftime("%H:%M:%S"))
#B12:
B12_jp2 = rasterio.open(ruta_B12)
nombre = str(ruta_B12)[76:91]
ruta_mediaB12 = carpetaintermedia + '\\\' + nombre + '_mediaB12.xlsx' #Se
guarda la media en unidades de reflectancia x 10000
#mediaParcela(B12_jp2, ruta_mediaB12)
df_B12 = pd.read_excel(ruta_mediaB12)
df_B12media = df_B12['media']/10000
B12_np = df_B12media.to_numpy()
print('Media B12 calculada')
print (time.strftime("%d/%m/%y"),',', time.strftime("%H:%M:%S"))

#Definición de las funciones para las dos opciones de cálculo:
##OPCIÓN A) Calcular primero el índice y luego la media por parcela (el
índice se calcula previo a utilizar la función)
def opcionA(INDICE_A, nombretiff, nombrexlsx, nombrecol):
    global df_INDICE_A
    #Convertir a raster:
    ruta_INDICE_A_raster = carpetaintermedia + '\\\' + nombre + '_' +
nombretiff
    guardarRaster(INDICE_A, ruta_INDICE_A_raster)
    #Calcular la media del índice por parcela:
    INDICE_tif = rasterio.open(ruta_INDICE_A_raster)
    ruta_mediaINDICE = carpetaintermedia + '\\\' + nombre + '_' +
nombrexlsx
    mediaParcela(INDICE_tif, ruta_mediaINDICE)
    df_INDICE_A = pd.read_excel(ruta_mediaINDICE)
    df_INDICE_A.columns = ['fid', 'cultivo', 'nº pixels', nombrecol]

#OPCIÓN B) Calcular los índices a partir de las medias de parcela (el
índice se calcula previo a utilizar la función)
def opcionB(INDICE_B, nombrecol):
    global df_INDICE_B
    df_INDICE_B = pd.DataFrame(INDICE_B)
    listafid = df_B4['fid']
    df_INDICE_B.insert(loc = 0, column = 'fid', value = listafid)
    df_INDICE_B.columns = ['fid', nombrecol]

#UNIÓN OPCIÓN A Y B:
def unionAB(df_INDICE_A, df_INDICE_B, INDICE, nombrexlsx):
    global df_INDICE#, df_corr

```

```

df_INDICE = df_INDICE_A.merge(df_INDICE_B, left_on = 'fid', right_on =
'fid')

df_corr = df_INDICE.corr()
ruta_INDICE = directorio + '\\ ' + nombre + '_' + nombrexlsx
writer = pd.ExcelWriter(ruta_INDICE, engine = 'xlsxwriter')
df_INDICE.to_excel(writer, sheet_name = INDICE + '_AB')
df_corr.to_excel(writer, sheet_name = 'Correlación')
writer.save()
writer.close()

##### CÁLCULO DE ÍNDICES #####
#### B8 Y B4
#### NDVI ####
def NDVI(Datos_B8, Datos_B4):
    return (Datos_B8 - Datos_B4) / (Datos_B8 + Datos_B4)
NDVI_A = NDVI(Datos_B8, Datos_B4)
NDVI_B = NDVI(B8_np, B4_np)
opcionA(NDVI_A, 'NDVI_A.tiff', 'NDVI_A.xlsx', 'NDVI_A')
opcionB(NDVI_B, 'NDVI_B')
unionAB(df_INDICE_A, df_INDICE_B, 'NDVI', 'NDVI.xlsx')
print('Índice NDVI calculado.')

#### SAVI ####
def SAVI(Datos_B8, Datos_B4, L):
    return ((Datos_B8 - Datos_B4) / (Datos_B8 + Datos_B4 * L)) * (1 + L)
SAVI_A = SAVI(Datos_B8, Datos_B4, 0.5)
SAVI_B = SAVI(B8_np, B4_np, 0.5)
opcionA(SAVI_A, 'SAVI_A.tiff', 'SAVI_A.xlsx', 'SAVI_A')
opcionB(SAVI_B, 'SAVI_B')
unionAB(df_INDICE_A, df_INDICE_B, 'SAVI', 'SAVI.xlsx')
print('Índice SAVI calculado.')

#### RDVI ####
def RDVI(Datos_B8, Datos_B4):
    return (Datos_B8 - Datos_B4) / ((Datos_B8 + Datos_B4)**(0.5))
RDVI_A = RDVI(Datos_B8, Datos_B4)
RDVI_B = RDVI(B8_np, B4_np)
opcionA(RDVI_A, 'RDVI_A.tiff', 'RDVI_A.xlsx', 'RDVI_A')
opcionB(RDVI_B, 'RDVI_B')
unionAB(df_INDICE_A, df_INDICE_B, 'RDVI', 'RDVI.xlsx')
print('Índice RDVI calculado.')

#### B12, B8 Y B4:
#### NDMI:
def NDMI(Datos_B8, Datos_B12):
    return (Datos_B8 - Datos_B12) / (Datos_B8 + Datos_B12)
NDMI_A = NDMI(Datos_B8, Datos_B12)
NDMI_B = NDMI(B8_np, B12_np)
opcionA(NDMI_A, 'NDMI_A.tiff', 'NDMI_A.xlsx', 'NDMI_A')
opcionB(NDMI_B, 'NDMI_B')
unionAB(df_INDICE_A, df_INDICE_B, 'NDMI', 'NDMI.xlsx')
print('Índice NDMI calculado.')

#### SLAVI ####
def SLAVI(Datos_B8, Datos_B12, Datos_B4):

```



```

        return (Datos_B8)/(Datos_B12 + Datos_B4)
    SLAVI_A = SLAVI(Datos_B8, Datos_B12, Datos_B4)
    SLAVI_B = SLAVI(B8_np, B12_np, B4_np)
    opcionA(SLAVI_A, 'SLAVI_A.tiff', 'SLAVI_A.xlsx', 'SLAVI_A')
    opcionB(SLAVI_B, 'SLAVI_B')
    unionAB(df_INDICE_A, df_INDICE_B, 'SLAVI', 'SLAVI.xlsx')
    print('Índice SLAVI calculado.')

#### B5, B6 Y B7:
#### CI-RE ####
def CIRE(Datos_B7, Datos_B5):
    return (Datos_B7/Datos_B5) - 1
CIRE_A = CIRE(Datos_B7, Datos_B5)
CIRE_B = CIRE(B7_np, B5_np)
opcionA(CIRE_A, 'CI-RE_A.tiff', 'CI-RE_A.xlsx', 'CI-RE_A')
opcionB(CIRE_B, 'CI-RE_B')
unionAB(df_INDICE_A, df_INDICE_B, 'CI-RE', 'CI-RE.xlsx')
print('Índice CI-RE calculado.')

#### MCARI ####
def MCARI(Datos_B5, Datos_B4, Datos_B3):
    return ((Datos_B5 - Datos_B4) - 0.2 * (Datos_B5 - Datos_B3)) *
(Datos_B5/Datos_B4)
MCARI_A = MCARI(Datos_B5, Datos_B4, Datos_B3)
MCARI_B = MCARI(B5_np, B4_np, B3_np)
opcionA(MCARI_A, 'MCARI_A.tiff', 'MCARI_A.xlsx', 'MCARI_A')
opcionB(MCARI_B, 'MCARI_B')
unionAB(df_INDICE_A, df_INDICE_B, 'MCARI', 'MCARI.xlsx')
print('Índice MCARI calculado.')

#### MTCI ####
MTCI_A = np.zeros(shape=(Datos_B5.shape[0], Datos_B5.shape[1]))
for i in range(Datos_B5.shape[0]):
    for j in range(Datos_B5.shape[1]):
        valor = Datos_B5[i][j] - Datos_B4[i][j]
        if valor == 0:
            MTCI_A[i][j] = 'nan'
        elif valor < 0 and valor > -0.1:
            MTCI_A[i][j] = (Datos_B6[i][j] -
Datos_B5[i][j])/(Datos_B5[i][j] - Datos_B4[i][j]-0.1)
        elif valor > 0 and valor < 0.1:
            MTCI_A[i][j] = (Datos_B6[i][j] -
Datos_B5[i][j])/(Datos_B5[i][j] - Datos_B4[i][j]+0.1)
        else:
            MTCI_A[i][j] = (Datos_B6[i][j] -
Datos_B5[i][j])/(Datos_B5[i][j] - Datos_B4[i][j])
    MTCI_B = np.zeros(shape=B6_np.shape[0])
    for i in range(B6_np.shape[0]):
        valor = B5_np[i] - B4_np[i]
        if valor == 0:
            MTCI_B[i] = 'nan'
        elif valor < 0 and valor > -0.1:
            MTCI_B[i] = (B6_np[i] - B5_np[i])/(B5_np[i] - B4_np[i] -0.1)
        elif valor > 0 and valor < 0.1:
            MTCI_B[i] = (B6_np[i] - B5_np[i])/(B5_np[i] - B4_np[i] +0.1)

```

```
else:
    MTCI_B[i] = (B6_np[i] - B5_np[i]) / (B5_np[i] - B4_np[i])

opcionA(MTCI_A, 'MTCI_A.tiff', 'MTCI_A.xlsx', 'MTCI_A')
opcionB(MTCI_B, 'MTCI_B')
unionAB(df_INDICE_A, df_INDICE_B, 'MTCI', 'MTCI.xlsx')
print('Índice MTCI calculado.')

#### NDVIre ####
def NDVIre(Datos_B7, Datos_B6):
    return (Datos_B7 - Datos_B6) / (Datos_B7 + Datos_B6)
NDVIre_A = NDVIre(Datos_B7, Datos_B6)
NDVIre_B = NDVIre(B7_np, B6_np)
opcionA(NDVIre_A, 'NDVIre_A.tiff', 'NDVIre_A.xlsx', 'NDVIre_A')
opcionB(NDVIre_B, 'NDVIre_B')
unionAB(df_INDICE_A, df_INDICE_B, 'NDVIre', 'NDVIre.xlsx')
print('Índice NDVIre calculado.')

#### SeLI ####
def SeLI(Datos_B8, Datos_B5):
    return (Datos_B8 - Datos_B5) / (Datos_B8 + Datos_B5)
SeLI_A = SeLI(Datos_B8, Datos_B5)
SeLI_B = SeLI(B8_np, B5_np)
opcionA(SeLI_A, 'SeLI_A.tiff', 'SeLI_A.xlsx', 'SeLI_A')
opcionB(SeLI_B, 'SeLI_B')
unionAB(df_INDICE_A, df_INDICE_B, 'SeLI', 'SeLI.xlsx')
print('Índice SeLI calculado.')

print('Carpeta', carpetas, 'procesada.')
print (time.strftime("%d/%m/%y"), ',', time.strftime("%H:%M:%S"))
```

#### Script 4. Interpolación de datos de Sentinel-1 a paso diario

```
import os
import pandas as pd
import glob
import numpy as np
import matplotlib.pyplot as plt

#Necesario instalar: openpyxl (desde sat2pyenv.bat: python -m pip install -U openpyxl)

df_ST = pd.read_excel(r'C:\DATA\TFM\Datos\ST_fechas.xlsx')
df_ST.fecha = df_ST.fecha.astype(str)
dir0 = r'C:\DATA\TFM\Datos\S1_r\Cultivos'
cont = os.listdir(dir0)

for pasada in cont:
    directorio = dir0 + '\\' + pasada
    contenido = os.listdir(directorio)
    for cultivo in contenido:
        dircultivo = directorio + '\\' + cultivo
        dircultivoxlsx = dircultivo + '\\*.xlsx'
        rutacultxlsx = glob.glob(dircultivoxlsx)
        for ruta in rutacultxlsx:
            print(ruta)
            df0 = pd.read_excel(ruta, header = 0)
            pos = ruta.rfind('_')
            IV = ruta[pos+1:-5]
            IV_A = IV + '_A'
            df = df0[['fecha', IV_A]]
            df.columns = ['fecha', IV]
            df_fecha = df['fecha']
            for i in range(len(df)):
                if len(str(df['fecha'].loc[i])) == 8:
                    df['fecha'].loc[i] = str(df['fecha'].loc[i])[:4] + '-' +
str(df['fecha'].loc[i])[4:6] + '-' + str(df['fecha'].loc[i])[6:]
                else:
                    df['fecha'].loc[i] = str(df['fecha'].loc[i])[:4] + '-' +
str(df['fecha'].loc[i])[4:6] + '-' + str(df['fecha'].loc[i])[6:8]
            listafechas = list(df['fecha'])
            df_g = df.groupby(['fecha']).mean()
            df_g = df_g.reset_index()
            df_n = df_ST.merge(df_g, how = 'outer')
            df_nn = df_n.fillna(0)

            inter = df_n.interpolate(method='linear',
axis=0,
limit=None,
inplace=False,
limit_direction='both',
limit_area=None,
downcast=None)

            dirfinal = r'C:\DATA\TFM\Datos\S1_r\Interpolados' + '\\' + pasada
            dirfinalxlsx = dirfinal + '\\' + cultivo + '\\Inter_' + cultivo +
 '_' + IV + '.xlsx'
            inter.to_excel(dirfinalxlsx, index = False)
```

**Script 5.** Aplicación del método Hold Out para separar el conjunto de test y entrenamiento.

```
from sklearn.model_selection import train_test_split
import numpy as np
import pandas as pd
import shutil
import os
import glob

cultivos = ['CEBADA SECANO', 'CEBADA REGADIO', 'GIRASOL', 'MAÍZ', 'TRIGO BLANDO SECANO', 'TRIGO BLANDO REGADIO', 'VIÑEDO VINIFICACIÓN']

np.random.seed(12)

dirorigen = r'C:\DATA\TFM\Datos\RELLENOHUECOS\1-DO'
dirtest = r'C:\DATA\TFM\Datos\RELLENOHUECOS\2-TEST'
dirtrain = r'C:\DATA\TFM\Datos\RELLENOHUECOS\2-TRAIN'

df =
pd.read_excel(r'C:\DATA\TFM\Datos\CULTIVOSINTERÉS\FidParcela_interes.xlsx')

for cultivo in cultivos:
    print(cultivo)
    if cultivo == 'CEBADA SECANO':
        Pcultivo = 'CEBADA'
        dfi = df.loc[(df['PRD_DESC'] == Pcultivo) & (df['COEF_REGAD'] == 0)]
    elif cultivo == 'CEBADA REGADIO':
        Pcultivo = 'CEBADA'
        dfi = df.loc[(df['PRD_DESC'] == Pcultivo) & (df['COEF_REGAD'] == 100)]
    elif cultivo == 'TRIGO BLANDO REGADIO':
        Pcultivo = 'TRIGO BLANDO'
        dfi = df.loc[(df['PRD_DESC'] == Pcultivo) & (df['COEF_REGAD'] == 100)]
    elif cultivo == 'TRIGO BLANDO SECANO':
        Pcultivo = 'TRIGO BLANDO'
        dfi = df.loc[(df['PRD_DESC'] == Pcultivo) & (df['COEF_REGAD'] == 0)]
    else: dfi = df.loc[df['PRD_DESC'] == cultivo]

    train, test = train_test_split(dfi, test_size = 0.50)
    lfidtrain = list(train['fid'])
    lfidtest = list(test['fid'])

    df_train = pd.DataFrame(lfidtrain)
    df_train.to_excel(os.path.join(dirtrain, 'Fidtrain_'+cultivo+'.xlsx'))
    df_test = pd.DataFrame(lfidtest)
    df_test.to_excel(os.path.join(dirtest, 'Fidtest_'+cultivo+'.xlsx'))

    print('train')
    for fidtrain in lfidtrain:
        inicitrain = glob.glob(os.path.join(dirorigen, cultivo, '*' +
str(fidtrain) + '*.xlsx'))
        fin0 = inicitrain[0].replace('RELLENOHUECOS\\1-DO\\', 'RELLENOHUECOS\\2-TRAIN\\')
        fin1 = inicitrain[1].replace('RELLENOHUECOS\\1-DO\\', 'RELLENOHUECOS\\2-TRAIN\\')
        shutil.copy(inicitrain[0], fin0)
        shutil.copy(inicitrain[1], fin1)
```

```
print('test')
for fidtest in lfidtest:
    iniciotest = glob.glob(os.path.join(dirorigen,cultivo,'*' +
str(fidtest) + '*.xlsx'))
    fin0 = iniciotest[0].replace('RELLENOHUECOS\\1-
DO\\','RELLENOHUECOS\\2-TEST\\')
    fin1 = iniciotest[1].replace('RELLENOHUECOS\\1-
DO\\','RELLENOHUECOS\\2-TEST\\')
    shutil.copy(iniciotest[0], fin0)
    shutil.copy(iniciotest[1], fin1)
```

**Script 6. Generador de huecos en las series temporales interpoladas del CT.**

```
import os
import glob
import pandas as pd
import numpy as np

df_ST = pd.read_excel(r'C:\DATA\TFM\Datos\ST_fechas.xlsx')
df_ST.fecha = df_ST.fecha.astype(str)
listafechas = list(df_ST['fecha'])

ventanaTI = ['2019-02-03', '2019-08-17'] #Invierno
ventanaTG = ['2019-04-04', '2019-10-16'] #Guisante
ventanaTM = ['2019-07-03', '2019-12-31'] #Maíz
ventanaTP = ['2018-09-01', '2019-12-31'] #Permanente: COGER TODO EL AÑO!!

VTIi = listafechas.index(ventanaTI[0])
VTIf = listafechas.index(ventanaTI[1])
VTGi = listafechas.index(ventanaTG[0])
VTGf = listafechas.index(ventanaTG[1])
VTMi = listafechas.index(ventanaTM[0])
VTMf = listafechas.index(ventanaTM[1])
VTPi = listafechas.index(ventanaTP[0])
VTPf = listafechas.index(ventanaTP[1])

lcultivos = ['CEBADA REGADIO', 'CEBADA SECANO', 'GIRASOL', 'MAÍZ', 'TRIGO BLANDO
REGADIO', 'TRIGO BLANDO SECANO', 'VIÑEDO VINIFICACIÓN']
lIVS2 = ['MCARI', 'MTCI', 'NDMI', 'NDVI', 'NDVIre']

diroriginal = r'C:\DATA\TFM\Datos\RELLENOHUECOS\3-TRAIN_INTER' #Partimos de
datos interpolados originales, NO SUAVIZADOS!!
dirfinal = r'C:\DATA\TFM\Datos\RELLENOHUECOS\3-TRAIN_INTER_HUECOS'

for cultivo in lcultivos:
    print(cultivo)
    VT = [VTIi, VTIf] if 'CEBADA' in cultivo or 'TRIGO' in cultivo else [VTMi,
VTMf] if 'MAÍZ' in cultivo else [VTGi, VTGf] if 'GIRASOL' in cultivo else
[VTPi, VTPf]
    tipoVentana = ventanaTI if 'CEBADA' in cultivo or 'TRIGO' in cultivo else
ventanaTM if 'MAÍZ' in cultivo else ventanaTG if 'GIRASOL' in cultivo else
ventanaTP
    dirS2cultivo = glob.glob(os.path.join(diroriginal,cultivo,'*S2*.xlsx'))
    for ruta in dirS2cultivo:
        #print(ruta)
```

```

df =
pd.read_excel(ruta).set_index('fecha')[tipoVentana[0]:tipoVentana[1]:5].reset_index()

for huecos in [1,2,3,4,5,10,15,20]: #AQUÍ ES DONDE SE VARÍA SI SE QUIEREN GENERAR HUECOS DIFERENTES.
    #Generar ST con huecos
    print(str(huecos) + ' huecos')
    for i in range(len(df)-(huecos-1)):
        #print(i)
        dfn = df.copy()
        listafechasalmacenadas = []
        for z in range(huecos):
            listafechasalmacenadas.append(dfn.iloc[i+z].fecha)
        dfn.set_index('fecha', inplace = True)
        for y in range(huecos):
            dfn.iloc[i+y] = np.nan
        pos0 = ruta.rfind('S2_')
        nombre = ruta[pos0+3:]
        dirfinalxlsx =
os.path.join(dirfinal,cultivo,str(huecos)+'d','STA_H'+str(huecos)+'d_S2_' +
listafechasalmacenadas[0]+'_'+listafechasalmacenadas[len(listafechasalmacenadas)-1] + '_' + nombre)
        dfn.to_excel(dirfinalxlsx,index=True)

```

#### Script 7. Técnica de relleno de huecos Chen et al (2004)

```

import os
import glob
import pandas as pd
import numpy as np
from scipy.signal import savgol_filter
import matplotlib.pyplot as plt
from sklearn.metrics import mean_squared_error

df_ST = pd.read_excel(r'C:\DATA\TFM\Datos\ST_fechas.xlsx')
df_ST.fecha = df_ST.fecha.astype(str)
listafechas = list(df_ST['fecha'])

#cultivos = ['CEBADA REGADIO', 'CEBADA SECANO', 'GIRASOL','MAÍZ','TRIGO BLANDO REGADIO','TRIGO BLANDO SECANO', 'VIÑEDO VINIFICACIÓN']
cultivos = ['MAÍZ']
dirTESTinter = r'C:\DATA\TFM\Datos\RELLENOHUECOS\7-TEST_VT_STA_HUECOS_INTER'
dirTESToriginal = r'C:\DATA\TFM\Datos\RELLENOHUECOS\5-TEST_VT_STA'
dirfinal = r'C:\DATA\TFM\Datos\RELLENOHUECOS\10-RELLENOHUECOS_NDVI\Apartir_S2'

carpetasdias = ['1d','2d','3d','4d','5d','10d','15d','20d']

IV = 'NDVI'

for cultivo in cultivos:
    print(cultivo)
    for dia in carpetasdias:
        print(dia)
        rmse0 = []

```



```

listaniter = []
listaniteroriginal = []
dircultivoTESTinterdia = os.path.join(dirTESTinter,cultivo,dia)
contenido = os.listdir(dircultivoTESTinterdia)
archivos =
len(os.listdir(os.path.join(dircultivoTESTinterdia,contenido[0])))
dffid = pd.DataFrame()
dffidoriginal = pd.DataFrame()
dfcultivo = pd.DataFrame()
for fid in contenido:
    posicion = contenido.index(fid)
    if posicion % 5 == 0:
print(str(posicion)+'/'+str(len(contenido))+'+'+str(dffid.shape))
    #print(fid)
    #print(dffid.shape)
    rutasTESTinter =
glob.glob(os.path.join(dircultivoTESTinterdia,fid,'*S2*.xlsx'))
    rutaTESTorigin =
glob.glob(os.path.join(dirTESTorigin,cultivo,'*S2*'+str(fid)+'*.xlsx'))
    dfTESTor = pd.read_excel(rutaTESTorigin[0]) [['fecha',IV]]
    dfTESTorig =
pd.DataFrame(pd.read_excel(rutasTESTinter[0]) ['fecha'])

##SOBRE LA SERIE ORIGINAL:
#print('Original')
dfTESToriginal = dfTESTorig.merge(dfTESTor)
# print('Huecos')
for rutaTESTinter in rutasTESTinter:
    #print(rutaTESTinter)
    pos0 = rutaTESTinter.rfind('_S2_')
    nom0 = rutaTESTinter[pos0+4:pos0+14]
    dfTEST = pd.read_excel(rutaTESTinter) [['fecha',IV]]
    x = dfTEST[IV].values
    df_SG = pd.DataFrame()
    df_SG[IV] = dfTEST[IV]
    xSG = savgol_filter(x, 19,1)
    df_SG[str(IV)+'_SG'] = xSG

#DETERMINAR EL PESO DE CADA PUNTO:
df_SG['abs'] = abs(df_SG[IV] - df_SG[str(IV)+'_SG'])
maxabs = df_SG['abs'].max()
df_SG['peso'] = np.where((df_SG[IV] >= df_SG[str(IV)+'_SG']),
1, (1-(df_SG['abs']/maxabs)))

#GENERAR NUEVA SERIE TEMPORAL DE NDVI:
##SOBRE LA SERIE DE HUECOS INTERPOLADA
i = 1
listaF = []
while i != 0:
    if len(listaF) > 1000: break
    #print(i)
    if i == 1:
        df_SG[str(IV)+str(i)] = np.where((df_SG[IV] >=
df_SG[str(IV)+'_SG']), df_SG[IV], df_SG[str(IV)+'_SG'])

```

```

        xfit = savgol_filter(df_SG[str(IV)+str(i)].values, 9,
6)

        df_SG[str(IV)+str(i)+'_fit'] = xfit
        #CALCULAR EL ÍNDICE DE EFECTO DE AJUSTE:
        df_SG['indiceEA_'+str(i)] =
abs(df_SG[str(IV)+str(i)+'_fit']-df_SG[IV])*df_SG['peso']
        Fanterior = df_SG['indiceEA_'+str(i)].sum()
        listaF.append(Fanterior)
        i +=1
    else:
        df_SG[str(IV)+str(i)] = np.where((df_SG[IV] >=
df_SG[str(IV)+str(i-1)+'_fit']), df_SG[IV], df_SG[str(IV)+str(i-1)+'_fit'])
        xfit = savgol_filter(df_SG[str(IV)+str(i)].values, 9,
6)

        df_SG[str(IV)+str(i)+'_fit'] = xfit
        #CALCULAR EL ÍNDICE DE EFECTO DE AJUSTE:
        df_SG['indiceEA_'+str(i)] =
abs(df_SG[str(IV)+str(i)+'_fit']-df_SG[IV]).values*df_SG['peso'].values
        Fsiguiente = df_SG['indiceEA_'+str(i)].sum()
        listaF.append(Fsiguiente)
        if i == 2:
            Factual = Fsiguiente
            i +=1
        else:
            if i >= 3:
                if Fanterior >= Factual and Factual <=
Fsiguiente:

                    niter = i-1
                    #print('El mejor ajuste es la iteración: '
+str(niter))

                    df_def = pd.DataFrame()
                    df_def[IV] =

df_SG[str(IV)+str(niter)+'_fit']

                    listaniter.append(niter)
                    dffid[nom0+'_'+str(fid)] = df_def[IV]
                    i = 0
                else:
                    Fanterior = Factual
                    Factual = Fsiguiente
                    i +=1

        #CÁLCULO DEL RMSE PARA CUANTIFICAR LA TÉCNICA DE RELLENO:
        #mse = mean_squared_error(
        #     y_true = dfTESToriginal[IV],
        #     y_pred = df_def[IV]
        # )
        #rmse = np.sqrt(mse)
        #rmse0.append(rmse)

        dffid_F = pd.concat([dfTESTorig,dffid], axis = 1)[:5]

dffid_F.to_excel(os.path.join(dirfinal,cultivo,'Predicciones_CHENetal_todo_'+d
ia+'_'+cultivo+'.xlsx'))
        rmselista = []
        fechaslistapasada = []

```

```
for col in list(dffid.columns):
    mse = mean_squared_error(
        y_true = dfTESToriginal[IV],
        y_pred = dffid[col]
    )
    rmse = np.sqrt(mse)
    rmselista.append(rmse)
    fechaslistapasada.append(col)
dfffinal = pd.DataFrame(list(zip(fechaslistapasada,rmselista)), columns
= ['fecha', 'RMSE'])

dfffinal.to_excel(os.path.join(dirfinal,cultivo,'RMSE_Predicciones_CHENetal_tod
o_'+dia+'_'+cultivo+'.xlsx'))
dfffinal[['fecha','fid']] = dfffinal.fecha.str.split(pat =
'_',expand=True)
dfffinal_group =
pd.DataFrame(dfffinal.groupby(['fecha']).mean()['RMSE']).reset_index()

dfffinal_group.to_excel(os.path.join(dirfinal,cultivo,'RMSE_Predicciones_CHENet
al_MEDIA_'+dia+'_'+cultivo+'.xlsx'))
dffid_T = dffid.T.reset_index()
columnas = ['fechafid' if x=='index' else x for x in
list(dffid_T.columns)]
dffid_T.columns = columnas
dffid_T[['fecha','fid']] = dffid_T.fechafid.str.split(pat =
'_',expand=True)
dff0 = dffid_T.groupby(['fecha']).mean()
dff_T = dff0.T
dfffinal = pd.concat([dfTESTorig,dff_T], axis = 1)[:5]

dfffinal.to_excel(os.path.join(dirfinal,cultivo,'Predicciones_CHENetal_MEDIA_'+
dia+'_'+cultivo+'.xlsx'))
```

### Script 8. Técnica de relleno de huecos a partir de datos Sentinel-1

```
import os
import glob
import pandas as pd
import numpy as np
from sklearn.metrics import mean_squared_error

df_ST = pd.read_excel(r'C:\DATA\TFM\Datos\ST_fechas.xlsx')
df_ST.fecha = df_ST.fecha.astype(str)
listafechas = list(df_ST['fecha'])

ventanaTI = ['2019-02-03', '2019-08-02'] #Invierno
ventanaTG = ['2019-04-04', '2019-10-21'] #Girasol
ventanaTM = ['2019-06-03', '2019-12-30'] #Maíz
ventanaTP = ['2018-09-01', '2019-12-30'] #Permanente: COGER TODO EL AÑO!!

VTIi = listafechas.index(ventanaTI[0])
VTIf = listafechas.index(ventanaTI[1])
VTGi = listafechas.index(ventanaTG[0])
VTGf = listafechas.index(ventanaTG[1])
VTMi = listafechas.index(ventanaTM[0])
VTMf = listafechas.index(ventanaTM[1])
VTPi = listafechas.index(ventanaTP[0])
VTPf = listafechas.index(ventanaTP[1])

#cultivos = ['CEBADA REGADIO', 'CEBADA SECANO', 'GIRASOL', 'MAÍZ', 'TRIGO BLANDO
REGADIO', 'TRIGO BLANDO SECANO', 'VIÑEDO VINIFICACIÓN']
cultivos = ['TRIGO BLANDO SECANO', 'VIÑEDO VINIFICACIÓN']

dirpredict0 = r'C:\DATA\TFM\Datos\RELLENOHUECOS\10-
RELLENOHUECOS_NDVI\Apartir_S1\OSERIES_REFERENCIA'
dirhuecos = r'C:\DATA\TFM\Datos\RELLENOHUECOS\6-TEST_VT_STA_HUECOS'
dirTESToriginal = r'C:\DATA\TFM\Datos\RELLENOHUECOS\5-TEST_VT_STA'
dirfinal = r'C:\DATA\TFM\Datos\RELLENOHUECOS\10-RELLENOHUECOS_NDVI\Apartir_S1'

LIVS1 = ['RDIF', 'RDIV', 'RMULTI', 'VH']
IVS2 = 'NDVI'

carpetasdias = ['1d', '2d', '3d', '4d', '5d', '10d', '15d', '20d']

for cultivo in cultivos:
    print(cultivo)
    dirpredict = os.path.join(dirpredict0, cultivo)
    VT = [VTIi, VTIf] if 'CEBADA' in cultivo or 'TRIGO' in cultivo else [VTMi,
VTMf] if 'MAÍZ' in cultivo else [VTGi, VTGf] if 'GIRASOL' in cultivo else
[VTPi, VTPf]
    tipoVentana = ventanaTI if 'CEBADA' in cultivo or 'TRIGO' in cultivo else
ventanaTM if 'MAÍZ' in cultivo else ventanaTG if 'GIRASOL' in cultivo else
ventanaTP
    id0 = listafechas.index(tipoVentana[0])
    id1 = listafechas.index(tipoVentana[1])
    fechas = listafechas[id0:id1]
    for IVS1 in LIVS1:
        print(IVS1)
```

```

for dia in carpetasdias:
    print(dia)
    dirhueco = os.path.join(dirhuecos,cultivo,dia)
    parcelas0 = os.listdir(dirhueco)
    parcelas = []
    for parcela in parcelas0:
        if len(parcela) < 8 and '65593' not in parcela:
            parcelas.append(parcela)
    dfmediasT_RL = pd.DataFrame()
    dfmediasT_RP2 = pd.DataFrame()
    dffids = pd.DataFrame()
    for fid in parcelas:
        posicion = parcelas.index(fid)
        if posicion % 5 == 0:
            print(str(posicion)+'/'+str(len(parcelas))+'+'+str(dffids.shape))
            #print(fid)
            dirhuecofid = glob.glob(os.path.join(dirhueco,fid,'*.xlsx'))
            dirpredictIVS1RL =
            glob.glob(os.path.join(dirpredict,'*RL*'+IVS1+'*.xlsx'))
            dirpredictIVS1RP2 =
            glob.glob(os.path.join(dirpredict,'*RP2*'+IVS1+'*.xlsx'))
            rutaTESTorigin =
            glob.glob(os.path.join(dirTESToriginal,cultivo,'*S2*'+str(fid)+'*.xlsx'))
            dfTESTor =
            pd.read_excel(rutaTESTorigin[0])[['fecha',IVS2]].set_index('fecha').T[fechas].
            T[:,5].reset_index()
            dfpredictRL0 = pd.read_excel(dirpredictIVS1RL[0])
            dfpredictRP20 = pd.read_excel(dirpredictIVS1RP2[0])
            listaindexRL = list(dfpredictRL0.index)
            listaindexRP2 = list(dfpredictRP20.index)
            posicionRL = listaindexRL.index(int(fid))
            dfpredictRL = dfpredictRL0.iloc[posicionRL].reset_index()
            dfpredictRL.columns = ['fecha',dfpredictRL.columns[1]]
            posicionRP2 = listaindexRP2.index(int(fid))
            dfpredictRP2 = dfpredictRP20.iloc[posicionRP2].reset_index()
            dfpredictRP2.columns = ['fecha',dfpredictRP2.columns[1]]
            df_reconstruida_RL = pd.DataFrame()
            df_reconstruida_RP2 = pd.DataFrame()
            j = 0
            for xlsx in dirhuecofid:
                # print(xlsx)
                pos0 = xlsx.rfind('S2_')
                fechai = xlsx[pos0+3:pos0+13]
                dfhueco = pd.read_excel(xlsx)[['fecha',IVS2]]
                dfhueco_RL = dfhueco.merge(dfpredictRL)
                dfhueco_RP2 = dfhueco.merge(dfpredictRP2)
                huecos_RL = dfhueco_RL[IVS2]
                huecos_RP2 = dfhueco_RP2[IVS2]
                predicciones_RL = dfhueco_RL[dfhueco_RL.columns[2]]
                predicciones_RP2 = dfhueco_RP2[dfhueco_RP2.columns[2]]
                serienueva_RL = []
                serienueva_RP2 = []
                for i in range(len(huecos_RL)):
                    serienueva_RL.append(predicciones_RL[i]) if
np.isnan(huecos_RL[i]) else serienueva_RL.append(huecos_RL[i])

```

```

        if j!=0:
            if len(serienueva_RL) == len(df_reconstruida_RL):
df_reconstruida_RL.insert(0,fechai,serienueva_RL)
            else: df_reconstruida_RL.insert(0,fechai,serienueva_RL)
            for i in range(len(huecos_RP2)):
                serienueva_RP2.append(predicciones_RP2[i]) if
np.isnan(huecos_RP2[i]) else serienueva_RP2.append(huecos_RP2[i])
            if j!=0:
                if len(serienueva_RP2) == len(df_reconstruida_RP2):
df_reconstruida_RP2.insert(0,fechai,serienueva_RP2)
                else: df_reconstruida_RP2.insert(0,fechai,serienueva_RP2)
                j +=1
            df_reconstruidaRL_F =
pd.concat([dfhueco['fecha'],df_reconstruida_RL], axis = 1)
            df_reconstruidaRL_fechas =
pd.DataFrame(dfTESTor['fecha']).merge(df_reconstruidaRL_F)

df_reconstruidaRL_fechas.to_excel(os.path.join(dirfinal,cultivo,'fid','Predicc
iones_RL_todos_'+str(fid)+'_'+IVS1+'_'+dia+'_'+IVS1+'_'+cultivo+'.xlsx'))
            df_reconstruidaRP2_F =
pd.concat([dfhueco['fecha'],df_reconstruida_RP2], axis = 1)
            df_reconstruidaRP2_fechas =
pd.DataFrame(dfTESTor['fecha']).merge(df_reconstruidaRP2_F)

df_reconstruidaRP2_fechas.to_excel(os.path.join(dirfinal,cultivo,'fid','Predic
ciones_RP2_todos_'+str(fid)+'_'+IVS1+'_'+dia+'_'+IVS1+'_'+cultivo+'.xlsx'))
            rmselistaRL = []
            rmselistaRP2 = []
            fechaslistapasada = []
            for col in list(df_reconstruida_RL.columns):
                mseRL = mean_squared_error(
                    y_true = dfTESTor[IVS2],
                    y_pred = df_reconstruidaRL_fechas[col]
                )
                rmseRL = np.sqrt(mseRL)
                mseRP2 = mean_squared_error(
                    y_true = dfTESTor[IVS2],
                    y_pred = df_reconstruidaRP2_fechas[col]
                )
                rmseRP2 = np.sqrt(mseRP2)
                rmselistaRL.append(rmseRL)
                rmselistaRP2.append(rmseRP2)
                fechaslistapasada.append(col)
            dffinal =
pd.DataFrame(list(zip(fechaslistapasada,rmselistaRL,rmselistaRP2)), columns =
['fecha','RMSE_RL','RMSE_RP2'])
            dffinal_fecha =
pd.DataFrame(dfTESTor['fecha']).merge(dffinal).set_index('fecha').T

dffinal_fecha.to_excel(os.path.join(dirfinal,cultivo,'fid','RMSE_Predicciones
_RL_RLP2_'+str(fid)+'_'+IVS1+'_'+dia+'_'+cultivo+'.xlsx'))
            colnuevas = [col+'_'+str(fid) for col in
dffinal_fecha.columns]
            dffinal_fecha.columns = colnuevas
            dffids = pd.concat([dffids,dffinal_fecha.T], axis = 0)

```



```
#dfreconstruidamedia_RL =
pd.DataFrame(df_reconstruida_RL.mean(axis = 1))
#dfreconstruidamedia_RL.columns = [fid]
#dfmediasT_RL =
pd.concat([dfmediasT_RL,dfreconstruidamedia_RL], axis = 1)
#dfreconstruidamedia_RP2 =
pd.DataFrame(df_reconstruida_RP2.mean(axis = 1))
#dfreconstruidamedia_RP2.columns = [fid]
#dfmediasT_RP2 =
pd.concat([dfmediasT_RP2,dfreconstruidamedia_RP2], axis = 1)
dffids_fechas = dffids.reset_index()
dffids_fechas.columns = ['fecha','RMSE_RL','RMSE_RP2']
dffids_fechas[['fecha','fid']] = dffids_fechas.fecha.str.split(pat
= '_',expand=True)
dffinal_group =
pd.DataFrame(dffids_fechas.groupby(['fecha']).mean()).reset_index()

dffinal_group.to_excel(os.path.join(dirfinal,cultivo,'RMSE_Predicciones_RL_RLP
2_MEDIA_'+IVS1+'_'+dia+'_'+cultivo+'.xlsx'))
```

**Script 9. Cálculo del RMSE para la evaluación de huecos.**

```
import os
import glob
import pandas as pd
from sklearn.metrics import mean_squared_error

df_ST = pd.read_excel(r'C:\DATA\TFM\Datos\ST_fechas.xlsx')
df_ST.fecha = df_ST.fecha.astype(str)
listafechas = list(df_ST['fecha'])

#lcultivos = ['CEBADA REGADIO', 'CEBADA SECANO', 'GIRASOL', 'MAÍZ', 'TRIGO
BLANDO REGADIO', 'TRIGO BLANDO SECANO', 'VIÑEDO VINIFICACIÓN']
lcultivos = ['MAÍZ']
carpetashuecos = ['1d', '2d', '3d', '4d', '5d', '10d', '15d', '20d']

dirSTA = r'C:\DATA\TFM\Datos\RELLENOHUECOS\5-TEST_VT_STA'
dirHUECO = r'C:\DATA\TFM\Datos\RELLENOHUECOS\7-TEST_VT_STA_HUECOS_INTER'
dirfidtest = r'C:\DATA\TFM\Datos\RELLENOHUECOS\2-TEST'
dirPasadas = r'C:\DATA\TFM\Datos\RELLENOHUECOS\0-PASADAS_CULTIVO'

ventanaTI = ['2019-02-03', '2019-08-17'] #Invierno
ventanaTG = ['2019-04-04', '2019-08-17'] #Girasol
ventanaTM = ['2019-06-03', '2019-12-31'] #Maíz
ventanaTP = ['2018-09-01', '2019-12-31'] #Permanente: COGER TODO EL AÑO!!

VTIi = listafechas.index(ventanaTI[0])
VTIf = listafechas.index(ventanaTI[1])
VTGi = listafechas.index(ventanaTG[0])
VTGf = listafechas.index(ventanaTG[1])
VTMi = listafechas.index(ventanaTM[0])
VTMf = listafechas.index(ventanaTM[1])
VTPi = listafechas.index(ventanaTP[0])
VTPf = listafechas.index(ventanaTP[1])

INDICE = 'NDVI'
dirfinal = r'C:\DATA\TFM\Datos\RELLENOHUECOS\8-RMSE_STAvsHUECOS_'+INDICE
sufijos = ['_HUECO', '_STA']

rxlsxfidtest0 = glob.glob(os.path.join(dirfidtest, '*.xlsx'))
rxlsxfidtest = []
for cultivo in lcultivos:
    for ruta in rxlsxfidtest0:
        if cultivo in ruta: rxlsxfidtest.append(ruta)
for i in range(len(rxlsxfidtest)):
    print(lcultivos[i])
    VT = [VTIi, VTIf] if 'CEBADA' in lcultivos[i] or 'TRIGO' in lcultivos[i]
else [VTMi, VTMf] if 'MAÍZ' in lcultivos[i] else [VTGi, VTGf] if 'GIRASOL' in
lcultivos[i] else [VTPi, VTPf]
    tipoVentana = ventanaTI if 'CEBADA' in lcultivos[i] or 'TRIGO' in
lcultivos[i] else ventanaTM if 'MAÍZ' in lcultivos[i] else ventanaTG if
'GIRASOL' in lcultivos[i] else ventanaTP
    df_fid = pd.read_excel(rxlsxfidtest[i]) #Para extraer los fid de cada
parcela de test
    listafid = list(df_fid[0])
    dircultivoSTA = os.path.join(dirSTA, lcultivos[i])
```

```

    dircultivoPasada =
glob.glob(os.path.join(dirPasadas, '*' + lcultivos[i] + '*.xlsx'))
    df_pasada = pd.read_excel(dircultivoPasada[0]) #Para extraer las fechas en
las que captura S2
    listapasada = list(df_pasada['fecha'])

    df_cultivo = pd.DataFrame(index = carpetashuecos, columns = listapasada)
#Df donde se irá guardando el RMSE de cada parcela

    for fid in listafid:
        #print(fid)
        dircultivoSTA =
glob.glob(os.path.join(dirSTA, lcultivos[i], '*S2*' + str(fid) + '*.xlsx'))
        dfSTA = pd.read_excel(dircultivoSTA[0]) [['fecha', INDICE]] if
len(dircultivoSTA) == 1 else print('MAL')
        df_cultivo_hueco = df_cultivo.copy()
        for hueco in carpetashuecos:
            dircultivoHUECO =
glob.glob(os.path.join(dirHUECO, lcultivos[i], hueco, str(fid), '*' + str(fid) + '*.xl
sx'))

            for ruta in dircultivoHUECO:
                #print(ruta)
                pos0 = ruta.rfind('_S2_')
                nom = ruta[pos0+4:]
                fechaini = nom[:10]
                fechafin = nom[11:21]
                if any(fecha in fechaini for fecha in listapasada):
                    dfHUECO = pd.read_excel(ruta) [['fecha', INDICE]]
                    #Cálculo del RMSE:
                    df0 = dfHUECO.merge(dfSTA, on = 'fecha',
suffixes=(sufijos[0], sufijos[1]))
                    RMSE = (mean_squared_error(df0[INDICE+sufijos[0]],
df0[INDICE+sufijos[1]]))**0.5

                    df_cultivo_hueco.loc[hueco][fechaini] = RMSE
                dircultivohuecofinal =
os.path.join(dirfinal, lcultivos[i], 'RMSE_' + INDICE + '_' + str(fid) + '_' + lcultivos[i]
+ '.xlsx')
                df_cultivo_hueco.to_excel(dircultivohuecofinal)

                if df_cultivo.isnull().all().all():
                    df_cultivo = df_cultivo_hueco
                else:
                    df_cultivo = (df_cultivo + df_cultivo_hueco) / 2
                dircultivofinal =
os.path.join(dirfinal, 'RMSE_medio_' + INDICE + '_' + lcultivos[i] + '.xlsx')
                df_cultivo.to_excel(dircultivofinal)

```